Journal of Experimental Psychology: General
2002, Vol. 131, No. 4, 552–566

0096-3445/02/$5.00 DOI: 10.1037//0096-3445.131.4.552

Constraints Have Different Concurrent Effects and Aftereffects on Variability

Patricia D. Stokes and Helen M. Harrison
Columbia University

Three experiments showed that constraints imposed early in learning have different effects on variability when they are in effect and after they are removed. Task constraints, which determine how something can be done, limited the number of possible responses in a computer game. Variability constraints, which specify how differently something must be done, required that each response differ from some number of prior responses. Less restrictive constraints (Experiments 1 and 2) produced higher variability during the constraints. More restrictive constraints (Experiments 2 and 3) led to higher variability after the constraints were relaxed. The authors discuss how these differences reflect strategies acquired during the constraints (default rules) and modified in closely related ways (exception rules) afterward.

Creativity has long been associated with variability (Csikszentmihalyi & Getzels, 1971; Gruber, 1988; Wallace & Gruber, 1989) but only recently with constraints (Stokes, 2001a, 2001b). In the arts, creative individuals maintain high levels of variability by means of self-imposed constraints. For example, to see how light breaks up on things, Monet first constrained light–dark contrasts, replacing them with contrasts between pure hues. To learn how light breaks up between things, he constrained motif. Painting the identical composition (e.g., the façade of Rouen cathedral, two haystacks in a field, a row of poplars reflected in the Seine) at different times of day, of year, and most importantly, in different light showed him, and us, what the *enveloppe* or atmosphere looks like.

Constraints imposed by a domain also promote variability. In poetry, traditional metric forms—alexandrine, hexameter, haiku—force writers to "search for words, select, reject, consider, make discoveries" (Byatt, 1991, p. 176). In architecture, multiple constraints—function, site, budget, buyer's taste—produce structures as different from each other as Frank Gehry's Guggenheimm Museum in Bilbao, Spain, with its shiny, splayed, industrial surfaces, and I. M. Pei's ethereal, pyramidal entrance for the Louvre in Paris.

This article does not argue that constraints per se produce breakthroughs. Rather, our interest is in how constraints—usually seen as restrictive and constrictive—can be used to increment the variability necessary for creativity, learning, and problem solving. We begin by reviewing current explanations of variable behavior.

Patricia D. Stokes and Helen M. Harrison, Department of Psychology, Columbia University.

This article is based on an undergraduate senior thesis submitted by Helen M. Harrison. We thank Allen Neuringer for suggestions that much improved our analyses.

Correspondence concerning this article should be addressed to Patricia D. Stokes, Department of Psychology, Barnard College, Columbia University, 3009 Broadway, New York, New York 10027. E-mail: pstokes@barnard.columbia.edu

Variability

Variability is defined as a continuum (Stokes, 1999b) with levels ranging from low to high. The levels are learned (Neuringer, 1993) and appear to be task or domain specific (Stokes, 2001–2002). An important issue concerns the content of that learning.

Neuringer (Grunow & Neuringer, 2002; Neuringer, 1993; Neuringer, Deiss, & Olson, 2000) was the first to show that variability can be selected and maintained by its consequences. In their classic study, Page and Neuringer (1985) required pigeons to produce sequences of eight pecks to right or left keys that differed from some number of immediately prior sequences. Variability increased as the number required increased. (The procedure is described more fully later in the section on *Variability Constraints.*) Page and Neuringer hypothesized that reward selects some output of an endogenous random variability generator. In this view, the selected variability level is the content of learning. Alternatively, Machado (1997) proposed that response patterns, which generate specific levels of variability, are the content of learning. Depending on the relative probabilities specified by a reinforcement schedule, some stable alternation pattern will earn maximal reward. Schedules that preclude strong sequential dependencies between responses generate high variability (Machado, 1992, 1994).

Our view is more eclectic. Taking an operant perspective, we have argued that learning involves how to do something and how differently to continue doing it—the latter is called a learned or habitual variability level (Stokes, 1995, 1999a; Stokes & Balsam, 2001). Supporting this view, variability levels acquired early in learning have been maintained in their respective training domains by rats (Stokes, 1995) and college students (Stokes & Balsam, 2001). We have also taken a cognitive, problem-solving perspective, suggesting that variability levels are incorporated into, or generated by, strategies acquired during learning (Stokes & Balsam, 2001; Stokes, Harrison, McElroy, & Paterniti, 1997; Stokes, Mechner, & Balsam, 1999).

A strategy is a response selected from a set of alternatives to reach a specific goal (Siegler, 1996; Wong, 1977). The set constitutes a hierarchy. The strategy with the most general condition and the greatest strength (due to past success) is the default rule. Other strategies in the hierarchy have more specific conditions and serve as exception rules to the default (Holland, Holyoak, Nisbett, & Thagard, 1987). Variability can follow from a strategy specifying a particular level or from a strategy generating a response pattern with its attendant level of variability.

Constraints

The problem-solving literature considers constraints two sided, as both limiting and directing search (Reitman, 1965). In other words, constraints preclude some responses to promote others. Some responses (e.g., reproduce exactly, give the single correct answer) indeed preclude variability and promote stereotypy. Those that increase variability do so by precluding reliable, predictable responses and promoting unusual, even novel ones (Stokes, 2001a, 2001b). Two kinds of constraints known to affect variability are task constraints, which determine how differently something can be done, and—given that multiple responses are possible in a given task—variability constraints, which specify how differently a task must be done.

Variability Constraints

Variability constraints specify either levels or kinds of variability. The first are quantitative, targeting how many different or same things are done. The second are qualitative, concerned, for example, with whether the things done are novel or fit in few or many categories.

Specific levels (high or low) can be selected by rewarding responses that differ from some number of prior responses. The required number is called a *lag*. For example, with a Lag 2 contingency on response sequences composed of four A and/or B responses, the third sequence in this series—AAAA, BBBB, AABB—would be rewarded, because it differs from the two previous ones. The third in this series—AABB, AAAA, AABB—would not be rewarded, because it repeats one of its two predecessors. Lag requirements have been used to increase variability in rats and pigeons (Hunziker, Saldana, & Neuringer, 1996; Morgan & Neuringer, 1990; Morris, 1987; Page & Neuringer, 1985), in college students (Wong & Peacock, 1986), and in children (Boulanger, 1990; Saldana & Neuringer, 1998).

When a lag is introduced determines whether high variability will be maintained after the constraint is removed. Studies with college students (Stokes, 1999a; Stokes & Balsam, 2001) have shown that early lag requirements continue to influence variability levels in the same task after the requirements are removed. Lags introduced later only increase variability when they are in effect.

An alternative procedure targets low-probability responses for reward (Bryant & Church, 1974; Machado, 1997; Maltzman, 1969; Neuringer, Kornell, & Oluffs, 2001). Machado (1994) called this frequency-dependent selection because high variability results from rewarding infrequent behaviors. For example, with a two-key apparatus, as responding on one key increases, the probability of reinforcement decreases on that key and increases on the other.

It should be noted that both frequency-dependent and lag procedures involve intermittent levels of reinforcement, which have been shown to increase (Tatham, Wanchisen, & Hineline, 1993), decrease (Herrnstein, 1961), or have inconsistent effects on (Eckerman & Lanson, 1969) variability. More relevant to the present study are the combined effects of variability constraints and reward levels. Looking at the relative contribution of each, Blough (1966) and Machado (1989) reported that variability was controlled by frequency-dependent requirements and not by rates of reinforcement. In an important recent study, Grunow and Neuringer (2002) demonstrated that the effects of intermittent reinforcement were inversely related to levels of variability prior to the decline in reward. When baseline variability was high, decreased reinforcement lowered it; when it was low, variability increased.

Kinds of variability include aspects of divergent thinking: (a) fluency, the number of different solutions; (b) flexibility, the number of solution categories (e.g., uses for a brick could fit in the categories "decoration," "weapon," or "building material"), and (c) originality, the number of unusual solutions (Guilford, 1950). Rewards (Campbell & Willis, 1978; Glover & Gary, 1976; Ryan & Winston, 1978), as well as instructions (Harrington, 1975; Runco, 1986), selectively increase the three types of divergent thinking. For example, Runco and Okuda (1991) showed that instructions to "be flexible, approach each question from different angles, and focus on variety" (p. 437) incremented (promoted) flexibility scores but depressed (precluded) those for originality.

Novelty is a kind of variability more difficult to promote because it is rare. Nonetheless, rewards have increased novelty (and precluded repetition) in the swimming and leaping topographies displayed by porpoises (Pyror, Haag, & O'Reilly, 1969), as well as in drawing and block-building forms produced by children (Funderbunk, 1977; Goetz, 1981; Holman, Goetz, & Baer, 1977). Here too, constraints continue to affect variability after they are relaxed. Holman et al. (1977) reported high levels of variability after reward for novelty ceased, as well as generalization within a training domain, for example, from drawing with felt pens to painting. More recently, Eisenberger and colleagues (Eisenberger & Armeli, 1997; Eisenberger & Selbst, 1994) demonstrated generalization across domains: Reward for novel drawings increased novelty in a subsequent, unrewarded unusual-uses task, and vice versa.

Task Constraints

Task constraints define domains and involve materials and conventions concerning their use. In the arts, creative individuals maintain high levels of variability through self-imposed task constraints, which determine how differently something can be done. For example, Jackson Pollack's so-called "drip" paintings were organized around a handful of purposefully placed gestures—elbow, X, and comma shapes. This gestural constraint provided the armature that supported a wide variety of paint applications: splashing, smudging, pouring, brushing, scrawling—with can, brush, and stick (Kimmelman, 1998). In other words, precluding variety in shape promoted variety in application. Both this and the earlier Monet examples involve task constraints imposed by the artists themselves (for other artistic examples, see Stokes, 2001a, 2001b). For the less creative among us, and certainly for experimental participants, task constraints are imposed by others.

One way in which all tasks constrain variability is in the number of alternatives they allow. A task popular with variability researchers involves making a series of left–right key presses to transverse a square or triangular matrix. We refer to these response sequences as *paths* and the places where they exit a matrix as *endpoints*. Some research using this task shows that having more alternatives increments variability. Increasing the number of response sequences from the upper-left to the lower-right corner of a square matrix from 70 (Schwartz, 1982a) to 256 (Page & Neuringer, 1985) increased variability. Page and Neuringer attributed the jump in variability to the laws of chance: Because the probability of repeating 1 of 70 responses (1/70) is higher than the probability of repeating 1 of 256 (1/256) responses, repetition should be higher and variability lower with fewer possible sequences. A triangular display with more paths (256 from apex to base) and more endpoints for the paths (nine) also increased variability relative to the original 70 path/one endpoint matrix (Tatham et al., 1993). Tatham et al. (1993) hypothesized that the additional exits may have "discouraged locking into a limited number of routes leading to a single exit point" (p. 358).

Studies with other tasks have shown that having fewer alternatives raises variability levels. When college students tried to produce useful inventions, those with no choice in either inventive category (e.g., furniture, toys) or parts to be used (e.g., wheels, handles) were more inventive than unconstrained students who were free to choose from one or both categories (Finke, 1990; Finke, Ward, & Smith, 1992). Compared with a group that could open a latch in any way, guinea pigs constrained to using a single paw were more variable in single-paw topography (Muenzinger, Koerner, & Irey, 1929). As with variability constraints, the timing of task constraints is important. A large early increase in the number of required responses led to higher sustained levels of variability in a computer game than exposure to the same shift only slightly later (Stokes et al., 1999). Likewise, early exposure to a more severe topographic constraint (press with right paw only vs. press any way) sustained more variable press topographies than introducing the same constraint later (Stokes, 1995).

Other task constraints known to affect variability involve the apparatus and the amount of effort involved. Using the same size and shape matrix as Schwartz (1982a), Wong and Peacock (1986) showed that replacing the original light panel with a video display increased variability, whereas increases in physical, temporal, or cognitive effort lowered it. For example, widely separated left (L) and right (R) response keys resulted in acquisition of a "block alternation strategy" (Wong & Peacock, 1986, p. 146) in which students pressed the maximum number of times on one key before switching to the other (e.g., RRRRLLLL, LLLLRRRR). Close placement of the left and right keys produced more within-sequence alternations (e.g., LRLRLRLR) and more different sequences.

The Present Experiments

According to this brief literature review, increased variability constraints increment variability, whereas increased task constraints cause either declines or increases. Our two primary goals were, first, to clarify how different task and variability constraints contribute to higher or lower learned variability levels and, second, to construe the contents of that learning.

A third goal was to examine the hypothesis that variability is the product of an innate quasi-random generator (Neuringer, 1986). We are cautious regarding this concept for a simple reason: Variability and randomness are not equivalent. Variability is a measure of how differently one behaves. Randomness is a measure of the predictability of that behavior. Nonrandom responding is predictable because it is structured; random responding is unpredictable because it is unstructured (Nickerson, 2002). Neither requires high variability nor precludes low variability. A Bach fugue is both highly structured and highly variable. The result of tossing a fair coin 25 times can be both unpredictable and invariant (e.g., all heads or all tails).

To these ends, we conducted three experiments to examine the effects of task and variability constraints on sustained or temporary levels of variability. We adapted a task used to show that more possibilities (lower task constraint) led to greater variability (Tatham et al., 1933). Our version involved two computer-generated triangles (hereinafter referred to as pyramids) that differed in number of possible paths and number of endpoints for those paths.

Experiment 1 compared responding during four different lags (variability constraints) on the two pyramids (task constraints). To test Page and Neuringer's (1985) law-of-chance hypothesis, we compared outputs of a random generator with student performance. Experiment 2 separated two aspects of the task constraint in Experiment 1, constraining either number of possible paths or endpoints on one size pyramid and examining variability during the constraint and after it was relaxed. Experiment 3 also looked at temporary and sustained variability levels, using separate and combined endpoint (task) and lag (variability) constraints.

Experiment 1: Combined Task and Variability Constraints

Experiment 1 compared responding on two versions of a computer game with a triangular display (the pyramid) in which each path from apex to base required 5 or 10 left and/or right arrow key presses. The pyramid requiring 10 presses had 1,024 possible paths and 11 endpoints; the one requiring 5 presses had 32 paths and 6 endpoints (task constraints). The same series of ascending lag requirements (variability constraints) was used on each pyramid. As discussed earlier, lag refers to the number of prior sequences from which a current sequence must differ.

If having fewer possibilities decreases variability, variability should be lower at each lag on the smaller than on the larger pyramid, and vice versa. If variability is due to students acting as if they were random generators, there should be no difference between the output of the students and the output of the random generator on the same size pyramids. If additional effort (5 more presses per path on the 10-pyramid) decreases variability, differences between the pyramids might be minimal.

Method

Participants

Ten first-year undergraduates at Barnard College, all of whom were women, participated to fulfill an introductory psychology class requirement.

Apparatus and Stimuli

Ten personal computers, in separate 1.5 m × 3.5 m experimental rooms, were used. Figure 1 shows what we refer to as a 5-pyramid. Superimposed on the pyramid was a grid with a white start box at the top and alternating red and blue diamonds between the white box and the bottom row of orange triangles, which were the endpoints. Pressing the left or right directional arrow keys moved the white box through the pyramid. A total of five presses, on right or left keys, moved the box from its starting position to one of the six orange triangles at the bottom.

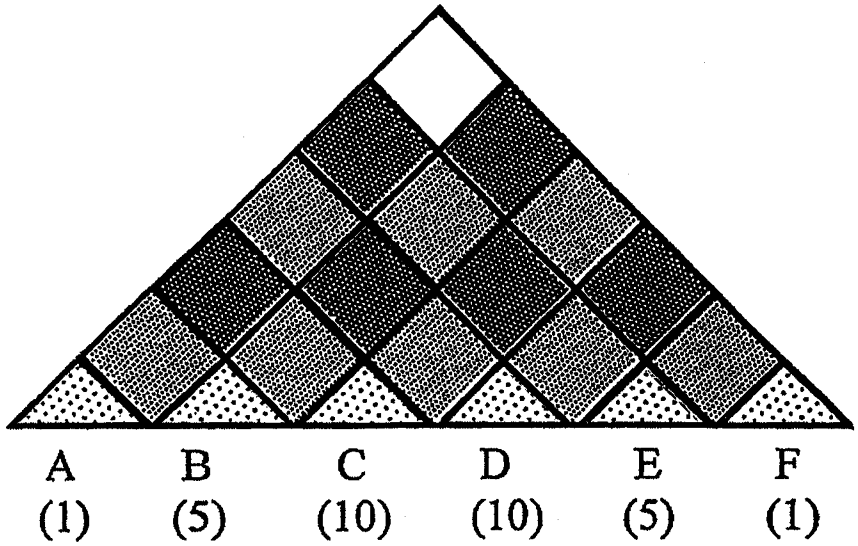

Figure 1. The 5-pyramid display with endpoints designated by letters A to F. Number of paths to each endpoint is shown in parentheses underneath.

When a press sequence met a contingency, the pyramid disappeared and the words "1 point" plus a cumulative total of points earned appeared, accompanied by a beep. At the end of each phase, the words "the session is over" appeared on the screen. Because 5 presses were required, this pyramid is called a 5-pyramid. The other pyramid (the 10-pyramid) required 10 presses.

Procedure

The following instructions were read to participants:

> The purpose of this experiment is to see how simple motor tasks are learned. You have been assigned to a group that gets feedback from the computer. Your task is to earn points by generating key press sequences. You can use two keys—the left and right directional arrows. After a correct sequence, a point will appear on the screen. During the experiment, the sequences that earn points may change. If this happens, you may have to alter your pressing patterns. You may not hold keys down for long periods of time. You must press a key and let it go. When the program stops, please come and get me.

Because we have been challenged for including the instruction alerting students that they may have to change the way they press, two demonstration groups (with and without this instruction) were run on an identical series of constraints. The results, which show no difference in outcome, are presented in the Appendix.

All of the participants worked on both pyramids. To avoid order effects, half of the participants worked for 100 reinforced trials on the 5-pyramid followed by 100 trials on the 10-pyramid, whereas the other half worked on the pyramids in reverse order. The sequence of lag requirements was identical on both pyramids. Lags increased from 0 to 2 to 10 to 20. During Lag 0, a current path did not have to differ from any prior path. Because the contingency reset at the start of each block, the Lag 2, 10, and 20 contingencies were based only on sequences emitted during those blocks and increased trial by trial until the full lag could be met. For example, during the Lag 2 block of trials, the number of prior paths from which a current path had to differ increased from 0 on the first trial to 1 on the second trial to 2 on all trials after the second trial. This meant that the first trial in the Lag 2 condition was always reinforced. The second trial was reinforced only if it differed from the first. After the third trial, each path had to differ from the previous two to earn points. A similar one-step-at-a-time increase in lag requirement occurred at the start of the Lag 10 and Lag 20 blocks. Each of the four lags lasted until 25 points were earned for correct sequences. There was no time limitation for completion of the 100 trials in each of the pyramids. Participants were debriefed at the end of the experiment. Ten separate outputs of a program generating random sequences of 5 or 10 lefts or rights were entered as data into the same sequences of lags on both sizes of pyramids.

Results

Student Data

To compare responding between pyramids and between lags, we ran mixed two-way analyses of variance (ANOVAs) with pyramid size (5 and 10) and lag (0, 2, 10, and 20) as variables.

Number of different paths. This is our measure of variability. The top panel of Figure 2 shows the mean number of different paths taken through the two pyramids at each lag. There were main effects of pyramid size, $F(1, 18) = 10.252$, $p < .01$, $\eta^2 = .363$, and lag, $F(3, 16) = 19.117$, $p < .001$, $\eta^2 = .782$. Number of different paths increased as the lag requirements increased on both pyramids and was greater at each lag on the larger pyramid. The interaction was not significant ($p = .45$).

Reinforcement percentages. Because reinforcement density may be related to variability, we calculated the percentage of reinforced paths during each lag. These are shown in the middle panel of Figure 2. There were significant main effects of pyramid size, $F(1, 18) = 22.509$, $p < .000$, $\eta^2 = .556$, and lag, $F(3, 16) = 54.834$, $p < .000$, $\eta^2 = .911$. As lags increased, percentages decreased on both pyramids, but more so on the smaller one. The interaction was significant, $F(3, 16) = 10.654$, $p < .000$, $\eta^2 = .666$. Paired sample t tests showed no changes between Lags 2 and 10 ($p = .951$) or between Lags 10 and 20 ($p = .225$) on the larger pyramid. On the smaller pyramid, percentages declined significantly between Lags 2 and 10, $t(9) = 2.939$, $p < .05$; the decline approached significance between Lags 10 and 20 ($p = .063$).

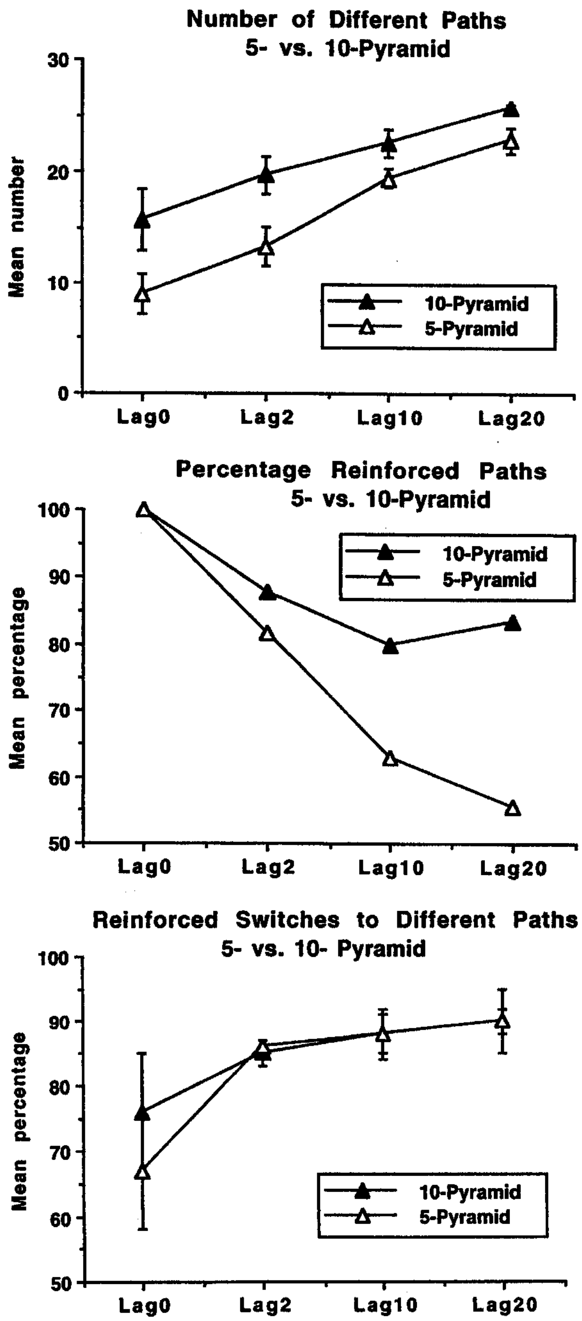

Figure 2. The three panels present data for responding on the 5-pyramid (white triangles) and the 10-pyramid (black triangles) during all lags: (top) mean number of different paths; (middle) mean percentage of reinforced paths; (bottom) mean percentage of reinforced switches between different paths.

To see if switching between paths was rewarded (and perhaps the source of high variability), we examined percentage of reinforcers earned for switching between different paths during each lag. These are shown in the bottom panel of Figure 2. There was no main effect of either lag ($p = .104$) or pyramid size ($p = .438$) and no significant interaction ($p = .438$). Because the mean percentages at Lag 0 do appear lower than the others, we compared Lags 0 and 2. Again, there were no significant effects: lag, $p = .063$; pyramid size, $p = .582$; interaction, $p = .459$.

Correlations. To see if either reinforcement density or reinforcement for switching was related to variability, we ran a Pearson correlation between number of different paths, percentage of reinforced paths, and percentage of reinforcers earned for switching between paths. Data were collapsed over both pyramids. Number of different paths and reinforcement for switching between paths were significantly correlated during Lag 0 ($r = .636$, $p < .05$). Number of different paths and percentage of reinforced paths were significantly correlated during Lag 20 ($r = .597$, $p < .005$). This last correlation is positive, indicating that higher levels of reinforcement, not lower ones, were associated with higher variability.

Number of presses. This is our measure of effort. Mean number of presses at each lag on the 5-pyramid were as follows: Lag 0 = 125, Lag 2 = 155, Lag 10 = 213, and Lag 20 = 239. On the 10-pyramid, the mean number of presses were as follows: Lag 0 = 250, Lag 2 = 293, Lag 10 = 300, and Lag 20 = 303. There were significant main effects of lag, $F(3, 16) = 19.628$, $p < .001$, $\eta^2 = .786$, and pyramid size, $F(1, 18) = 59.575$, $p < .001$, $\eta^2 = .768$. Number of presses increased as lags increased and was higher at all lags on the larger pyramid. The interaction was significant, $F(3, 16) = 3.172$, $p = .05$, $\eta^2 = .373$. Paired sample t tests showed that number of presses increased significantly between Lags 0 and 2, $t(9) = 2.479$, $p < .05$, and between Lags 2 and 10, $t(9) = -3.408$, $p < .01$, on the smaller pyramid, but only between Lags 0 and 2, $t(9) = 3.224$, $p < .01$, on the larger.

Self-reports. During debriefing, students were asked what they did to earn points. Their descriptions were sorted into three separate and two combined categories. All nonspecific references to varying went into the *vary* category. Table 1 shows the results. Fifty percent of the students simply said they went to endpoints at the base of the pyramid, 30% mentioned pattern making, 10% said they used different paths to an endpoint, and another 10% (in the vary category) said they acted randomly. Only 2 students distinguished between the two size pyramids; both said that it was harder to earn points on the smaller one.

Table 1
Percentage of Students' Self-Reports

Category	Patterns	Both[a]	Endpoints	Both[a]	Vary
		Experiment 1			
5- & 10-pyramids	30	10	50		10
		Experiment 2			
1 endpoint	60		10	20	10
2 endpoints (10 paths)	40		50		10
4 endpoints	80		0		20
20 paths	60		30		10
		Experiment 3			
Lag	47		13		60
Endpoint	47	13	33		7
Combo	13	67	7		13

[a] *Both* refers to a combination of adjacent categories, not a distinction between them.

Student Versus Random Sequence Data

To see if higher variability on the 10-pyramid was due to students acting like random number generators, we entered 10 separate outputs of the random right–left sequence generator (equal to the number of students) as responses on the computer game for both pyramids and compared them with student performance. To equalize number of paths, we compared the last 25 trials during each lag for the student and the random data. Separate paired-sample t tests were run for data on each pyramid. The top panel of Figure 3 shows number of different paths taken during the last 25 trials on the 5-pyramid.

The students differed from the random generator at Lag 0, $t(18) = -3.967, p < .001$; Lag 2, $t(18) = -3.208, p < .01$; and Lag 10, $t(18) = -2.250, p < .05$. Students and random generator did not differ at Lag 20 ($p = .165$). The bottom panel of Figure 3 shows number of different paths on the 10-pyramid. The students differed from the random generator at all lags: Lag 0, $t(18) = -3.333, p < .01$; Lag 2, $t(18) = -4.157, p < .001$; Lag 10, $t(18) = -3.613, p < .01$; and Lag 20, $t(18) = -5.532, p < .001$.

These results suggest that responding was more systematic or structured on the 10-pyramid than on the 5-pyramid. One metric indicative of structure, derived from information theory (Miller & Frick, 1949), is an uncertainty. Uncertainty is a function of possible outcomes. Our measure, next-press uncertainty, reflects the probability of a right press being followed by a right or left press, and vice versa. As the internal structure of a path becomes less orderly, the uncertainty will increase. Conversely, the more stereotyped the internal structure, the lower the uncertainty will become. Next-press uncertainty was computed using the formula,

$$U[R/(R-1)] = U(R, R-1) - U(R-1),$$

where $U[R/(R-1)]$ is the uncertainty in the joint distribution of the current (R) and immediately prior (R − 1) responses. Figure 4 shows next-press uncertainties for all responses on the two pyramids during each lag. A two-way ANOVA with pyramid size (5, 10) and lag (0, 2, 10, 20) as variables produced main effects of pyramid size, $F(1, 18) = 5.134, p < .05, \eta^2 = .222$, and lag, $F(3, 16) = 8.473, p < .001, \eta^2 = .614$. The interaction was not significant ($p = .769$). Next-press uncertainty was higher at every lag on the 5-pyramid, confirming that responding was more orderly on the 10-pyramid.

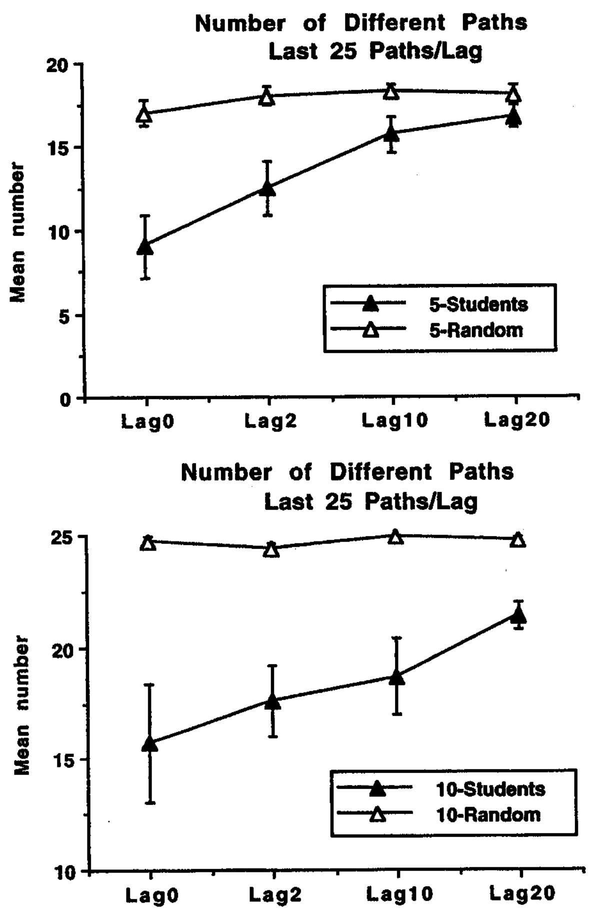

Figure 3. Mean number of different paths during the last 25 trials of each lag for students (black triangles) and random generated (white triangles) data: (top) data from the 5-pyramid; (bottom) data from the 10-pyramid.

Discussion

Our main question concerned the effects of task constraints. These differed in number of possible paths and endpoints, as well as in effort required.

Did more possibilities increase variability? Yes. Number of different paths taken was greater at all lags on the pyramid with more paths and endpoints. Did greater effort decrease variability? No. Although number of presses was higher at each lag on the larger compared with the smaller pyramid, variability was higher on the larger pyramid. Did students perform like the random number generator? Only on the smaller pyramid and only at the highest lag, never on the larger pyramid.

Our second question involved the contents of learning. The uncertainty analysis indicated that responding was more structured on the larger pyramid. To help explain this greater orderliness, we examined the raw data. The most noticeable pattern involved wedge-shaped blocks in which the number of right or left presses decreased or increased by one in each successive trial. In the example shown in Figure 5, the wedge shapes are separated. Two variants of this pattern were observed. In one, the decreasing series of Rs or Ls was followed by alternating Ls and Rs (e.g., RRRRL, RRRLR, RRLRL, RLRLR). In the other, it was followed by a single change of press (e.g., RRRRL, RRRLR, RRLRR, RLRRR).

The wedge pattern is interesting for several reasons. First, it could have been easily generated by a visual algorithm or strategy, for example, "go straight down the right side of the pyramid, go straight down the right side and turn left 1 up from the end . . . 2 up . . . 3 up . . . and so on." Second, it may explain why, in the debriefing, most students reported aiming at endpoints. Systematically shifting one left or right press in successive trials produces paths that sweep (in an very orderly manner) across the bottom of the pyramid in one direction and then the other. Third, patterns of these kinds would produce many more different paths on the 10-pyramid (11 in one direction, 22 if continued in the reverse direction) than on the 5-pyramid (6 in one direction, 12 if reversed).

To see if this were the case, we counted number of successive paths in each lag block that represented this kind of patterning, dividing by the total number of paths in the block to get percentage of paths in patterned blocks. At all lags, the wedge pattern ap-

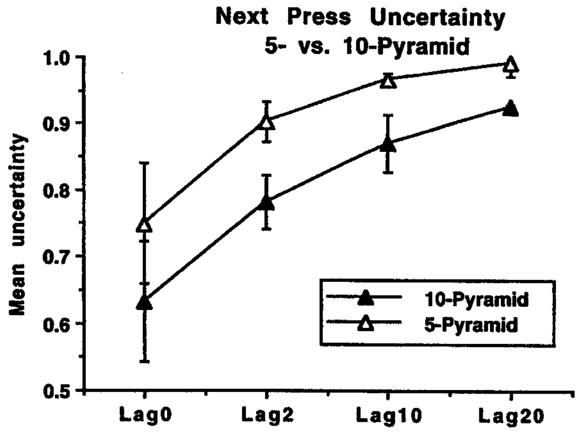

Figure 4. Next-press uncertainty for all trials of each lag on the 5-pyramid (white triangles) and 10-pyramid (black triangles).

peared more often on the larger pyramid, in which the percentages were as follows: Lag 0, 42%; Lag 2, 50%; Lag 10, 46%; and Lag 20, 30%. On the smaller pyramid, they were as follows: Lag 0, 27%; Lag 2, 39%; Lag 10, 21%; and Lag 20, 14%.

Although the wedge clearly cannot be the sole structuring element on the larger pyramid, it does lend support to Tatham et al.'s (1993) hypothesis that more exits (endpoints) increase variability. The pattern is not possible on a square matrix with only one exit, which is the kind used by Schwartz (1982a, 1982b).

In sum, we are left with Tatham et al.'s (1993) hypothesis: More paths and more exits (endpoints) increase variability. The design of Experiment 1 did not let us separate the effects of possible paths and possible endpoints, nor did it let us look at responding after the constraints were relaxed. Experiment 2 was designed to do both.

Experiment 2: Task Constraints

Given the results of Experiment 1, number of paths or number of endpoints affects variability. In Experiment 2, using one size pyramid allowed either number of endpoints or paths to be varied in different groups. It is important to note that while it was possible

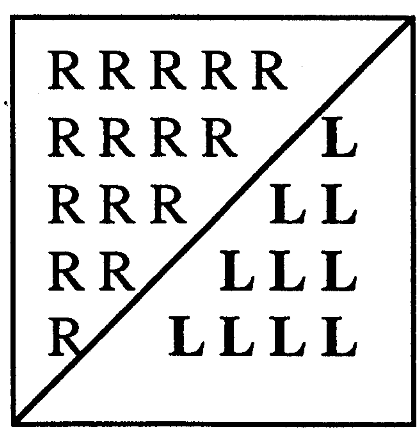

Figure 5. Example of a wedge-pressing pattern in which the number of right (R) presses decreases in each successive path, whereas the number of left (L) presses increases.

to vary within each set of task constraints, there was no explicit variability constraint: A single path could meet all requirements. To see if either or both task constraints had lasting effects on variability, we looked at responding during the constraints and after they were relaxed.

Method

Participants

Forty Barnard College women undergraduates participated to fulfill an introductory psychology class requirement.

Apparatus and Stimuli

The apparatus was identical to that used in Experiment 1. The stimulus was the 5-pyramid. As shown in Figure 1, this pyramid has six endpoints, identified in the figure by letters A to F. The number of unique paths leading to each endpoint is also shown.

Procedure

Ten participants were randomly assigned to each of four constrained groups. Table 2 shows the conditions.

The structure of the pyramid made it impossible to hold number of paths completely constant while varying the number of endpoints across a reasonable range. However, number of paths in the endpoint constraint groups was quite close: 10 (with 1 or 2 endpoints) or 12 (with 4 endpoints). In the group with 1 endpoint, 5 participants were rewarded for exiting at the C endpoint; the other 5 were rewarded for exiting at the D endpoint. Data from the group labeled B&E (10) in Table 2 were included in both endpoint (2) and path (10) comparisons. The instructions read were the same as those used in Experiment 1. The constraints were in effect until 50 points were earned. One hundred additional points were earned after the constraints were relaxed.

Results

The experiment lasted for 150 reinforced trials, which—for analyses—were divided into blocks of 25 reinforced trials each. The last block in the constraint phase (26th to 50th reinforced trials) and the last in the free phase (126th to 150th) were examined. Mixed two-way ANOVAs with group (1 endpoint, 2 endpoints [10 paths], 4 endpoints, 20 paths) and phase (constraint or free) as variables were used to compare responding between these blocks. Fisher's least significant difference (LSD) test was used for post hoc comparisons. A significance level of .05 was used.

Two sets of comparisons were made: between groups in which either paths (10–12) were relatively constant and number of end-

Table 2
Path and Endpoint Constraints in Experiment 2[a]

	Path	
Endpoint	10–12	20
1	C or D (10)	
2	B&E (10)[b]	C&D
4	A&B&E&F (12)	

[a] Refer to Figure 1 for the paths (in parentheses) and endpoints (Letters A to F). [b] Data from this group were included in both endpoint (2) and path (10) comparisons.

points (1, 2, or 4) varied or endpoints were held constant (2) and number of paths varied (10–20).

Number of Different Paths

Figure 6 presents our measure of variability: mean number of different paths for the endpoint constraint groups during the constraint and free phases. Figure 7 shows data for the path constraint groups. The top panel shows results for the constraint phase; the bottom shows results for the free phase. There were significant main effects of phase, $F(1, 36) = 14.407, p < .001, \eta^2 = .286$, and group, $F(3, 36) = 3.174$, $p < .05$, $\eta^2 = .209$. Because the interaction was not significant ($p = .113$), paired-sample t tests were run. These showed that number of different paths increased significantly between the constraint and free phases in groups 1 endpoint, $t(9) = -3.278$, $p < .01$, and 2 endpoints (10 paths), $t(9) = -2.765, p < .05$, but not in 4 endpoints ($p = .343$) or 20 paths ($p = .256$).

The effects of endpoints and paths were evaluated with LSD tests. As the bottom panel of Figure 6 shows, number of early endpoints was inversely related to variability during the free phase: The 1 endpoint group was significantly more variable than the 4 endpoints group ($p < .05$). The top panel of Figure 7 shows that number of paths was directly related to variability during the constraint phase: The group with 20 paths was more variable than the 10-paths group ($p < .05$).

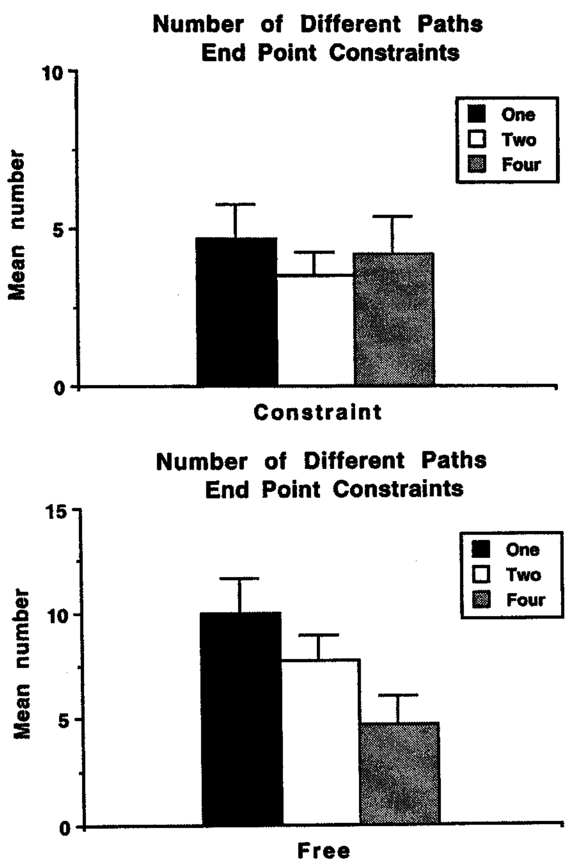

Figure 6. Mean number of different paths taken by the endpoint constraint groups during the constraint (top) and free (bottom) phases of Experiment 2.

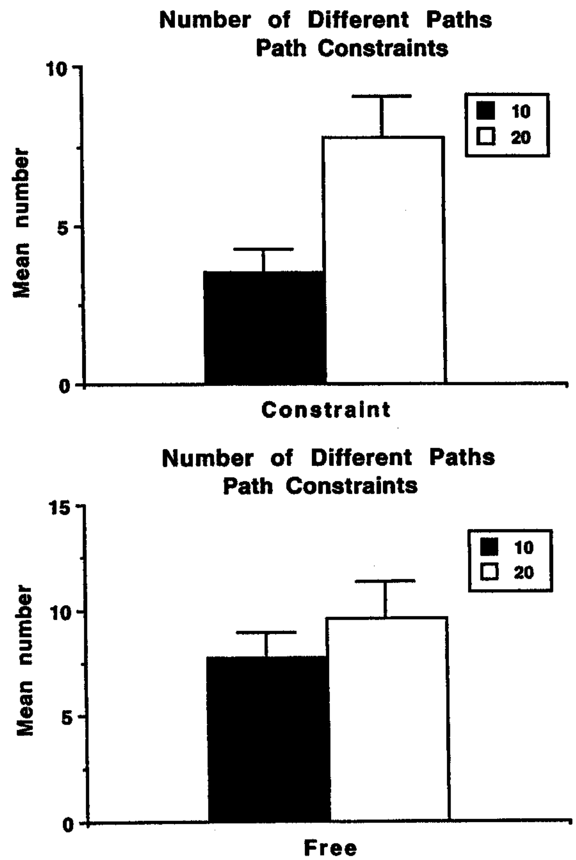

Figure 7. Mean number of different paths taken by the path constraint groups during the constraint (top) and free (bottom) phases of Experiment 2.

Reinforcement Percentages

To determine if reinforcement densities differed, we examined percentages of reinforced paths. During the constraint phase, these were as follows: for endpoint constraint, the 1 endpoint group, 96.7%; the 2 endpoints group, 95.2%; and the 4 endpoints group, 97.1%; for path constraint, the 10-paths group, 95.2% and the 20-paths group, 96.9%. During the free phase, all paths were reinforced (100%). There was a significant main effect of phase, $F(1, 36) = 33.043, p < .000, \eta^2 = .479$. Reinforcement increased between the constraint and free phases for all groups. There was no effect of group ($p = .686$) and no interaction of phase and group ($p = .686$).

To see if switching was rewarded, we examined percentage of points earned for switching between different paths. Figure 8 shows these percentages for all lags in the endpoint constraint group. Figure 9 shows them for the path constraint groups.

There were main effects of phase, $F(1, 36) = 18.390, p < .001$, $\eta^2 = .338$, and group, $F(3, 36) = 2.835, p = .05, \eta^2 = .191$. The top panel of Figure 8 shows that during the constraint phase, the 2 endpoints group earned fewer points for switching than the 4 endpoints group ($p < .05$). The bottom panel shows that during the

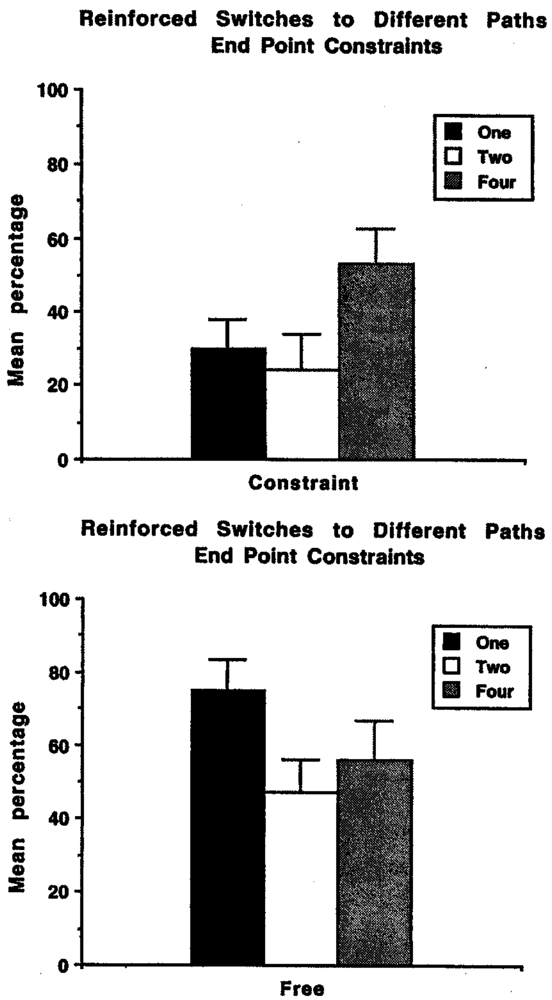

Figure 8. Mean percentage of reinforced switches to different paths for the endpoint constraint groups during the constraint (top) and free (bottom) phases of Experiment 2.

free phase, the 1 endpoint group earned more than the 2 endpoints group ($p < .05$). Figure 9 shows that during the (top panel) constraint ($p < .05$) and (bottom panel) free phases ($p < .05$), the 20-paths group earned more points for switching than the 10-paths group. The Phase × Group interaction was also significant, $F(3, 36) = 3.484$, $p < .05$, $\eta^2 = .225$. Increases in the free phase were greater in the 1 than the 2 or 4 endpoints groups and in the 10- versus the 20-paths group.

Correlations

To see if reinforcement density or reward for switching was related to variability, we collapsed data from the two phases. A Pearson correlation was run between number of different paths, percentage of reinforced paths, and percentage of reinforcers earned for switching between paths. During both the constraint ($r = .594$, $p < .001$) and free ($r = .586$, $p < .001$) phases, number of different paths was positively correlated with percentage of reinforcers earned for switching between paths.

Number of Different Endpoints

Number of different endpoints was analyzed to see if participants came in contact with the relaxed constraints during the free phase. The mixed ANOVA showed a main effect of phase, $F(1, 36) = 34.399$, $p < .01$, $\eta^2 = .489$, but not of group. There was also a significant Group × Phase interaction, $F(3, 36) = 4.240$, $p < .05$, $\eta^2 = .261$.

Number of endpoints increased between the constraint ($M = 2.53$) and free ($M = 4.42$) phases in all groups, indicating that they did contact the change in contingency. The increases were greatest in the groups with 1 or 2 initial endpoints (constraint = 2.3, free = 4.4) and least in the group with 4 initial endpoints (constraint = 3.1, free = 3.4). LSD tests showed that during the constraint phase, the 4 endpoints group went to more endpoints than the 1 endpoint group ($p = .011$).

Self-Reports

Students' descriptions of what they did to earn points were sorted into the categories shown in Table 1, which presents the percentages for each group.

As the totals show, the majority in three groups mentioned patterns. The patterns described varied by group: The 1 endpoint group specified zig-zagging, making diagonals, and systematic alternations; in the less variable 4 endpoints group, 3 students said

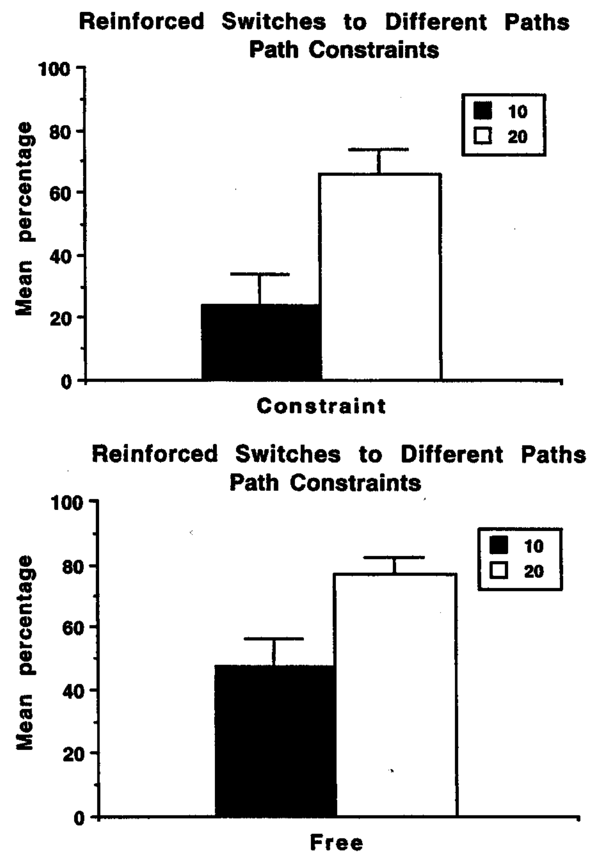

Figure 9. Mean percentage of reinforced switches to different paths for the path constraint groups during the constraint (top) and free (bottom) phases of Experiment 2.

they went straight down the sides, and 2 others stuck to one pattern. Mentions of endpoints referred to the specific number restricted initially in the 1 and 2 endpoints groups (e.g., one first, later all). The combination endpoint–vary students in the 1 endpoint group said they first went to one endpoint and then kept changing. Most in the vary category said they pressed randomly.

It is important to note that several students mentioned that they discovered the change in requirements by "accident" or "mistake." That is, while aiming for a particular endpoint, they exited at a different one and found that it too earned points.

Discussion

Our main question again concerned the effects of the task constraints. Experiment 1 indicated that either or both more paths and endpoints led to higher variability. Experiment 2 showed that path and endpoint constraints had different consequences when they were in effect and after they were relaxed. Less restrictive task constraints (more possible paths) led to more different paths taken during the constraint, but not later. More restrictive ones (fewer possible endpoints) led to more different paths taken after, but not during, the constraint.

Our second question concerned the contents of learning. The majority of students focused on patterns, indicating that either specific sequences (e.g., all left or right presses) or strategies (e.g., reverse or mirror a previous path) were learned. When endpoints were mentioned, it was in terms of changes in constraint (from one or two to all). As in Experiment 1, the least frequently mentioned alternative was varying per se. Another clue came from the Pearson correlations, showing that reward for switching between paths was positively correlated with variability (number of different paths) during both constraint and free phases. This suggests acquisition of a strategy (or a visual algorithm as in Experiment 1) involving systematic switching, which dovetails nicely with student reports of reversing paths, alternating directions, and so on.

The distribution of strategies differed from that in Experiment 1. When students were required to vary but were not restricted in endpoints (note that the 5-pyramid had fewer endpoints than the 10-pyramid, but no endpoint constraint was in effect on either pyramid), most focused on endpoints. In Experiment 2, when they were not required to vary but were restricted to specific endpoints, most focused on patterns. Obviously, we needed to compare the separate and combined effects of variability and endpoint constraints.

Experiment 3: Variability, Task, and Combined Constraints

Experiment 3 looked at temporary and sustained variability levels by using separate and combined task and variability constraints. There were two reasons for doing this.

The first was variability differences in the first two experiments. In Experiment 1, lag requirements and number of possible paths affected variability while each constraint was in effect. Likewise, in Experiment 2, possible paths only affected variability when the constraints were in effect. However, number of endpoints affected variability levels after the constraints were relaxed. Experiment 3 let us compare the current and sustained effects of each type of constraint and their combination.

The second was strategy differences. In Experiment 1, when specific levels of variability were required, the most-mentioned strategy involved endpoints. In Experiment 2, when specific endpoints were required, the predominant focus was on patterns. Experiment 3 would reveal what happens when both variability levels and endpoints are specified. To these ends, Experiment 3 used one display (the 5-pyramid), one variability constraint (a moderate lag), one task constraint (a single endpoint), and their combination.

Method

Participants

Forty-five Barnard and Columbia College undergraduates (49 women, 6 men) participated to fulfill an introductory psychology class requirement.

Apparatus and Stimuli

Apparatus and stimuli were identical to those used in Experiment 2.

Procedure

The instructions used in Experiment 2 were read to all participants, who were randomly assigned to one of three constraint groups. The constraints were in effect until 50 points were earned. Group names indicate the type of constraint. In the lag group, only paths that differed from five prior paths were rewarded. In the endpoint group, reward followed paths exiting the pyramid at the one endpoint, C. In the combination group, reward depended on two things: exiting at the C endpoint and doing so taking a path that differed from five prior paths. Once 50 points were earned, the constraints were relaxed. The experiment lasted for 100 additional trials, during which any path earned points.

Results

The experiment lasted for 150 reinforced trials, which—for analyses—were divided into blocks of 25 reinforced trials each. As in Experiment 2, the last block in the constraint phase (26th to 50th reinforced trials) and in the free phase (126th to 150th reinforced trials) were examined. Mixed two-way ANOVAs with group (lag, endpoint, or combination) and phase (constraint or free) as variables were used to compare responding between these blocks. Fisher's LSD test was used for post hoc comparisons. A significance level of .05 was used.

Number of Different Paths

As in Experiments 1 and 2, this is our measure of variability. Figure 10 shows the mean number of different paths taken by each group during the (top panel) constraint and (bottom panel) free phases. The ANOVA showed a main effect of group, $F(2, 42) = 9.013$, $p < .001$, $\eta^2 = .300$, and a significant Phase × Group interaction, $F(2, 42) = 9.016$, $p < .001$, $\eta^2 = .300$. There was no main effect of phase ($p = .203$). Number of different paths increased between the constraint and free phases in the endpoint, decreased in the lag, and remained the same in the combination groups.

The LSD tests showed that during the constraint phase, the two groups with the lag requirement (lag and combination) took more different paths than the endpoint group (both $ps < .001$). During the free phase, the combination group became the most variable,

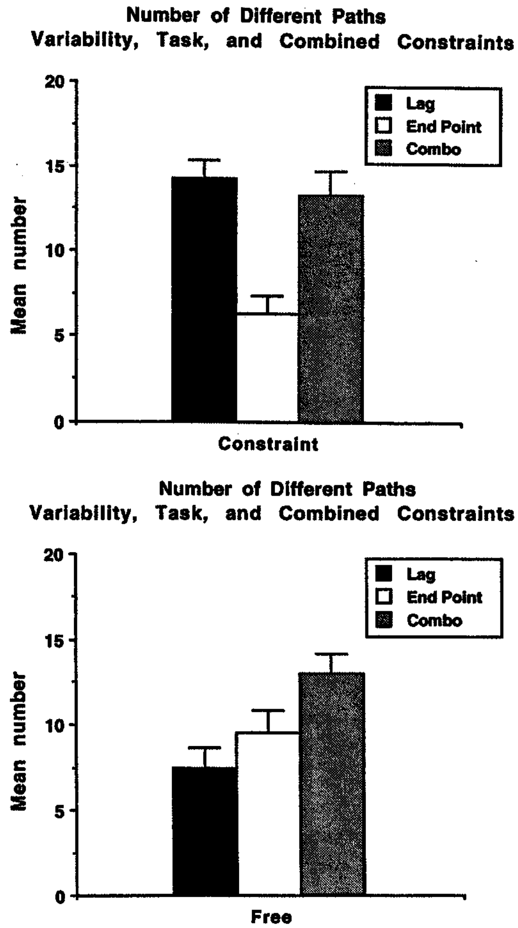

Figure 10. Mean number of different paths taken by the lag, endpoint, and combination groups during the constraint (top) and free (bottom) phases of Experiment 3.

taking more paths than either the endpoint ($p < .05$) or lag ($p < .01$) groups.

Reinforcement Percentages

To see if reinforcement densities differed, we examined percentage of reinforced paths. During the constraint phase, these were as follows: lag group, 88%; endpoint group, 93%; and combination group, 55%. In the free phase, 100% of paths were reinforced in all groups. There were significant main effects of phase, $F(1, 42) = 137.192$, $p < .001$, $\eta^2 = .766$, and group, $F(2, 42) = 38.552$, $p < .001$, $\eta^2 = .647$. The interaction was also significant, $F(2, 42) = 38.552$, $p < .001$, $\eta^2 = .647$. Reinforcement increased from the constraint to the free phases for all groups. The greatest increase was in the combination group. The LSD tests showed that, during the constraint phase, both the lag and endpoint groups had higher percentages of reinforced paths than the combination group (both $ps < .001$).

Figure 11 shows the percentage of reinforcers earned for switching between paths. The top panel presents results from the constraint phase; the bottom panel presents results from the free phase. There were main effects of group, $F(2, 42) = 8.199$, $p < .001$, $\eta^2 = .281$, and phase, $F(1, 42) = 17.791$, $p < .001$, $\eta^2 = .298$, as well as a significant Group × Phase interaction, $F(2, 42) = 15.659$, $p < .001$, $\eta^2 = .427$. LSD tests showed that during the constraint phase, the lag group was reinforced for switching more often than the endpoint or combination groups (both $ps < .001$). Although there were no significant differences during the free phase, the group that was reinforced most for switching (combination group) also took the highest number of different paths.

Correlations

To see if there were any relationships between variability (number of different paths) and reinforcement, we ran a Pearson correlation between number of different paths, percentage of reinforced paths, and percentage of reinforcers earned for switching between paths. Data were collapsed over groups. During the constraint phase, number of different paths was negatively correlated with percentage of reinforced paths ($r = -.418$, $p < .01$) and positively with percentage of reinforced switches to different paths ($r = .505$, $p < .001$). During the free phase, number of different paths was positively correlated with percentage of reinforced switches to different paths ($r = .650$, $p < .001$).

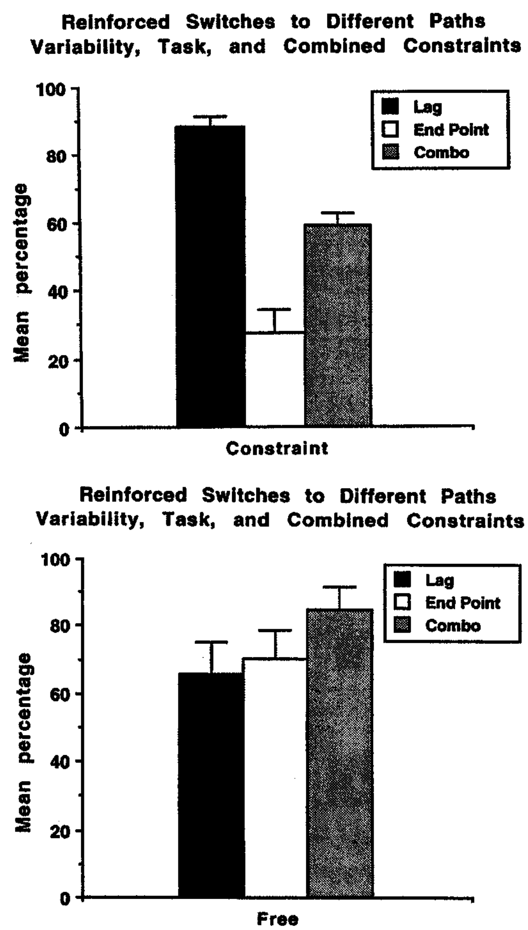

Figure 11. Mean percentage of reinforced switches to different paths taken by the lag, combination, and endpoint groups during the constraint (top) and free (bottom) phases of Experiment 3.

Our conclusion is a caveat. Be careful, not casual, about the constraints you choose to introduce new skills. They will determine how variably your students use those skills.

References

Blough, D. S. (1966). The reinforcement of least-frequent interresponse times. *Journal of the Experimental Analysis of Behavior, 9,* 581–591.

Boulanger, B. (1990). *La variabilite comportementale: Une approache developpementale chez l'humain* [Behavioral variability: A developmental approach]. Unpublished doctoral dissertation, Faculte de Psychologie et des Sciences de l'Education, Universite de Liege, Liege, Belgium.

Bryant, D., & Church, R. M. (1974). The determinants of random choice. *Animal Learning and Behavior, 2,* 245–248.

Byatt, A. S. (1991). The TLS poetry competition. In A. S. Byatt (Ed.), *Passions of the mind* (pp. 174–176). New York: Vintage International.

Campbell, J. A., & Willis, J. (1978). Modifying components of "creative behavior" in the natural environment. *Behavior Modification, 2,* 549–565.

Csikszentmihalyi, M., & Getzels, J. W. (1971). Discovery-oriented behavior and the originality of artistic products: A study with artists. *Journal of Personality and Social Psychology, 19,* 47–52.

Eckerman, D. A., & Lanson, R. N. (1969). Variability of response location for pigeons responding under continuous reinforcement, intermittent reinforcement, and extinction. *Journal of the Experimental Analysis of Behavior, 12,* 73–80.

Eisenberger, R. (1992). Learned industriousness. *Psychological Review, 99,* 248–267.

Eisenberger, R., & Armeli, S. (1997). Can salient reward increase creative performance without reducing intrinsic creative interest? *Journal of Personality and Social Psychology, 72,* 652–663.

Eisenberger, R., Masterson, F. A., & McDermitt, M. (1982). Effects of task variety on generalized effort. *Journal of Educational Psychology, 74,* 499–505.

Eisenberger, R., & Selbst, M. (1994). Does reward increase or decrease creativity? *Journal of Personality and Social Psychology, 66,* 1116–1127.

Finke, R. (1990). *Creative imagery: Discoveries and inventions in visualization.* Hillsdale, NJ: Erlbaum.

Finke, R., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications.* Cambridge, MA: MIT Press.

Funderbunk, F. R. (1977). Reinforcement control of classroom creativity. In T. Brigham, R. Hawkins, J. W. Scott, & R. F. McLaughlin (Eds.), *Behavior analysis in education: Self-control and reading* (pp. 197–204). Dubuque, IW: Kendall/Hunt.

Goetz, E. M. (1981). The effects of minimal praise on the creative blockbuilding of three year olds. *Child Study Journal, 11,* 55–67.

Glover, J. A., & Gary, A. L. (1976). Procedures to increase some aspects of creativity. *Journal of Applied Behavior Analysis, 9,* 79–84.

Gruber, H. E. (1988). The evolving systems approach to creative work. *Creativity Research Journal, 1,* 27–51.

Grunow, A., & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychonomic Bulletin and Review, 9,* 250–258.

Guilford, J. P. (1950). Creativity. *American Psychologist, 5,* 444–454.

Hackenberg, T. D., & Joker, V. R. (1994). Instructions versus schedule control of humans' choices in situations of diminishing returns. *Journal of the Experimental Analysis of Behavior, 62,* 367–383.

Harrington, D. M. (1975). Effects of explicit instructions to "be creative" on the psychological meaning of divergent thinking tests. *Journal of Personality, 43,* 434–454.

Herrnstein, R. J. (1961). Stereotypy and intermittent reinforcement. *Science, 133,* 2067–2069.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1987). *Induction: Processes of inference, learning, and discovery.* Cambridge, MA: MIT Press.

Holman, J., Goetz, E. M., & Baer, D. M. (1977). The training of creativity as an operant and an examination of its generalization characteristics. In B. C. Etzel, J. M. LeBlanc, & D. M. Baer (Eds.), *New developments in behavioral research: Theory, method and applications* (pp. 441–471). Hillsdale, NJ: Erlbaum.

Hunziker, M. H. L., Saldana, R., & Neuringer, A. (1996). Behavioral variability in shy and why rats as a function of rearing environment and reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 65,* 129–144.

Joyce, J. H., & Chase, P. N. (1990). Effects of response variability on the sensitivity of rule-governed behavior. *Journal of the Experimental Analysis of Behavior, 54,* 251–262.

Kimmelman, H. (1998, October 30). How even Pollack's failures enhance his triumph. *New York Times,* pp. E1, E3.

Lai, Q., & Shea, C. (1998). Generalized motor program (GMP) learning: Effects of reduced frequency of knowledge of results and practice variability. *Journal of Motor Behavior, 30,* 51–59.

Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedule. *Journal of the Experimental Analysis of Behavior, 42,* 155–166.

Machado, A. (1992). Behavioral variability and frequency-dependent selection. *Journal of the Experimental Analysis of Behavior, 58,* 241–263.

Machado, A. (1994). Polymorphic response patterns under frequency-dependent selection. *Animal Learning and Behavior, 22,* 53–71.

Machado, A. (1997). Increasing the variability of response sequences in pigeons by adjusting the frequency of switching between two keys. *Journal of the Experimental Analysis of Behavior, 68,* 1–25.

Maltzman, I. (1969). On the training of originality. *Psychological Review, 67,* 229–242.

Miller, G., & Frick, F. C. (1949). Statistical behavioristics and sequences of responses. *Psychological Review, 56,* 311–324.

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

Morris, C. J. (1987). The operant conditioning of response variability: Free-operant versus discrete-response procedures. *Journal of the Experimental Analysis of Behavior, 47,* 273–277.

Muenzinger, K., Koerner, L., & Irey, E. (1929). Variability of an habitual movement in guinea pigs. *Journal of Comparative Psychology, 9,* 425–436.

Neuringer, A. (1986). Can people behave "randomly"? The role of feedback. *Journal of Experimental Psychology: General, 115,* 62–75.

Neuringer, A. (1993). Reinforced variation and selection. *Animal Learning and Behavior, 21,* 83–91.

Neuringer, A., Deiss, C., & Olson, G. (2000). Reinforced variability and operant learning. *Journal of Experimental Psychology: Animal Behavior Processes, 26,* 98–111.

Neuringer, A., Kornell, N., & Oluffs, M. (2001). Effects of extinction on response variability and structure. *Journal of Experimental Psychology: Animal Behavior Processes, 1,* 79–94.

Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review, 2,* 330–357.

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

Pyror, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior, 12,* 653–661.

Reitman, W. R. (1965). *Cognition and thought: An information processing approach.* New York: Wiley.

Runco, M. A. (1986). Flexibility and originality in children's divergent thinking. *Journal of Psychology, 120,* 345–352.

Runco, M. A., & Okuda, S. M. (1991). The instructional enhancement of the flexibility and originality scores of divergent thinking tests. *Applied Cognitive Psychology, 5,* 435–441.

Ryan, B. A., & Winston, A. S. (1978). Dimensions of creativity in children's drawings: A social validation study. *Journal of Experimental Psychology, 70,* 651–656.

Saldana, R. L., & Neuringer, A. (1998). Is instrumental variability abnormally high in children exhibiting ADHD and aggressive behavior? *Behavior Brain Research, 94,* 51–59.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptions of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3,* 207–217.

Schwartz, B. (1982a). Failure to produce response variability with reinforcement. *Journal of the Experimental Analysis of Behavior, 37,* 171–181.

Schwartz. B. (1982b). Reinforcement-induced stereotypy: How not to teach people to discover rules. *Journal of Experimental Psychology: General, 111,* 23–59.

Siegler, R. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Simon, D. A., & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 907–912.

Stokes, P. D. (1995). Learned variability. *Animal Learning and Behavior, 23,* 164–176.

Stokes, P. D. (1999a). Learned variability levels: Implications for creativity. *Creativity Research Journal, 12,* 37–45.

Stokes, P. D. (1999b). Novelty. In M. A. Runco & S. R. Pritzker (Eds.), *Encyclopedia of creativity* (Vol. II, pp. 297–304). San Diego, CA: Academic Press.

Stokes, P. D. (2001a). Constraints and creativity: Experiments in art and psychology. In S. Nagel (Ed.), *The handbook of policy creativity: Vol. III. Creativity from diverse perspectives* (pp. 33–48). Huntington, NY: Nova Science.

Stokes, P. D. (2001b). Variability, constraints, and creativity: Shedding light on Claude Monet. *American Psychologist, 36,* 355–359.

Stokes, P. D. (2001–2002). Variations on Guilford's creative abilities. *Creativity Research Journal, 13,* 277–283.

Stokes, P. D., & Balsam, P. (1991). Effects of reinforcing preselected approximations on the topography of the rat's barpress. *Journal of the Experimental Analysis of Behavior, 12,* 349–373.

Stokes, P. D., & Balsam, P. (2001). An optimal period for setting sustained variability levels. *Psychonomic Bulletin and Review, 8,* 177–184.

Stokes, P. D., Harrison, H., McElroy, T., & Paterniti, A. (1997, March). *Effects of strategic hints on task variability.* Paper presented at the meeting of the Eastern Psychological Association, Washington, DC.

Stokes, P. D., Mechner, F., & Balsam, P. D (1999). Effects of different acquisition procedures on response variability. *Animal Learning and Behavior, 27,* 28–41.

Tatham, T. A., Wanchisen, B. A., & Hineline, P. N. (1993). Effects of fixed and variable ratios on human behavioral variability. *Journal of the Experimental Analysis of Behavior, 59,* 349–359.

Wallace, D. B., & Gruber, H. E. (1989). *Creative people at work: Twelve cognitive case studies.* New York: Oxford University Press.

Weiner, H. (1964). Conditioning history and human fixed-interval performance. *Journal of the Experimental Analysis of Behavior, 7,* 383–385.

Wong, P. T. (1977). Partial reinforcement and the learning of a strategy in the rat. *American Journal of Psychology, 90,* 663–674.

Wong, P. T., & Peacock, E. J. (1986). When does reinforcement induce stereotypy? A test of the differential reinforcement hypothesis. *Learning and Motivation, 17,* 139–161.