

How examples may (and may not) constrain creativity

RICHARD L. MARSH, JOSHUA D. LANDAU, and JASON L. HICKS
University of Georgia, Athens, Georgia

Three experiments were performed to test Smith, Ward, and Schumacher's (1993) conformity hypothesis—that people's ideas will conform to examples they are shown in a creative generation task. Conformity was observed in all three experiments; participants tended to incorporate critical features of experimenter-provided examples. However, examination of total output, elaborateness of design, and the noncritical features did not confirm that the conformity effect constrained creative output in any of the three experiments. Increasing the number of examples increased the conformity effect (Experiment 1). Examples that covaried features that are naturally uncorrelated in the real world led to a greater subjective rating of creativity (Experiment 2). A delay between presentation and test increased conformity (Experiment 3), just as models of inadvertent plagiarism would predict. The explanatory power of theoretical accounts such as activation, retrieval blocking, structured imagination, and category abstraction are evaluated.

Traditionally, creativity has been studied as an individual difference measured on tests such as the Remote Associates (Mednick, 1962) or the Torrance battery (Torrance, 1974), and sundry measures of divergent thinking (e.g., devising novel uses for a brick). Although valuable, that approach has tended to underemphasize the existence of common cognitive processes and representations that support creative thinking. Recently, however, Ward, Smith, Finke, and their colleagues (e.g., Finke, Ward, & Smith, 1992; Smith, Ward, & Schumacher, 1993; Ward, 1994) have made valuable empirical and theoretical strides by delineating some of the universal cognitive mechanisms that underlie people's production of new ideas and novel solutions to creative problems. One theoretical construct, *structured imagination*, proposes that when faced with a problem whose solution requires creativity, people may take a *path of least resistance* by retrieving domain-specific information or an existing solution and then attempting to modify that old construct in some novel way (Ward, 1994).

The retrieval or heightened accessibility in memory of existing (and perhaps suboptimal) solutions in the preceding account leads to documented regularities in people's creative output (see, e.g., Jansson & Smith, 1991; Smith et al., 1993; Ward, 1994). More specifically, people's creative solutions tend to incorporate features of ei-

ther examples they are shown or related exemplars that they are able to retrieve from memory. In brief, creative solutions are rarely truly novel, insofar as they usually contain vestiges of prior solutions to the problem. In the experiments that follow, Smith et al.'s (1993) and Ward's (1994) paradigms were adopted in order to assess the degree to which incorporating features of old solutions actually constrains people's creativity. Our goal is to isolate the locus of any such constraints by examining the specific impact that incorporating old features might have on both the novel and the not-so-novel aspects of people's creations.

Structured imagination is an approach whose goals are similar to those of the experiments reported here because of that theory's focus on both what is old and what is new in a novel creation (Ward, 1995). According to structured imagination, when creating something novel, people bring to mind (either consciously or unconsciously) existing categories and concepts, thereby causing the features of those categories to become incorporated into their creations. In four experiments, Ward (1994) asked people to generate non-Earth-like creatures to inhabit a new planet. People tended to adopt as their starting point the domain of Earth animals, which then led to creations that quite uniformly displayed the attributes of bilateral symmetry, sense organs for obtaining information, and appendages for locomotion and manipulation of the environment (as well as many other Earth-animal attributes). When constraints were placed on the uninhabited planet (e.g., molten lava), people tended to place wings on their creations for flight as opposed to including features from other, nonanimal categories such as jet engines or propellers (Ward, 1994, Experiment 3). Instructions to create creatures that were beyond people's wildest imagination and without regard to whether someone would believe they could actually exist did not diminish this regularity

This research was supported in part by a University of Georgia Research Foundation, Inc. grant awarded to the first author and a Sigma Xi Grant-in-Aid to each of the second and third authors. Appreciation is expressed to Evan Heit and to Tom Ward for insightful comments expressed on an earlier version. Special thanks go to Jonathan Reaves, D. J. Amis, and Christopher Marino for help in scoring the data. Correspondence regarding this article should be addressed to R. L. Marsh, Department of Psychology, University of Georgia, Athens, GA 30602-3013 (e-mail: marsh@meme.psy.uga.edu).

and the structured nature of their creative output (Experiment 4). Because there is remarkable agreement on typical attributes of category members (e.g., Tversky & Hemenway, 1984), relying on the category of "Earth animals" made people's novel creations rather predictable in terms of the features they contained. Therefore, imagination and creativity appear to be fairly structured (or constrained) by the cognitive processes and strategies that support them.

To the extent that retrieval of prior examples or domain information may limit creativity, making specific examples more available should bias people's creations toward the features contained in those examples. Smith et al. (1993) tested this prediction in three experiments by comparing example groups who received three experimenter-provided novel creations with control groups who did not. People in the example conditions tended to incorporate three critical features (four legs, tail, and antennae) contained in all of the experimenter-provided examples, whereas people in the control conditions incorporated them less frequently (the same was true of three critical features of novel toys). Importantly, this *conformity* effect appeared to be an unintentional use of memory, because people who were instructed to diverge as much as possible from the examples nevertheless incorporated the three critical features into their novel creations (Smith et al., 1993, Experiment 3). Taken together, these experiments and Ward's (1994) results suggest that regardless of the source (either experimenter-provided or self-generated), activated information can lead to conformity—which is a form of constraint on creativity.

These effects are conceptually related to the inadvertent plagiarism (technically termed *cryptomnesia*), found when people engage in creative problem solving (Brown & Murphy, 1989; Marsh & Bower, 1993). In those tasks, participants were asked to generate category exemplars or to solve word puzzles. In both cases, additional solutions were presented from another source (i.e., additional items were encountered other than what the participant generated). When asked to recall one's own solutions or to provide brand new items that had not previously been generated before, people tend to claim that solutions previously offered by others are actually their own new, novel contributions. And this inadvertent theft occurs despite strong admonitions not to repeat an earlier presented item. Marsh and Landau (1995) have argued that when people engage in creative tasks they fail to closely monitor the source of their solutions; and as a consequence, information that has retained activation from prior exposure is more likely to be adopted or incorporated into a task at hand (cf. Jacoby, Kelley, Brown, & Jasechko, 1989). This explanation is consistent with the failure to find a decrease in conformity when people are explicitly instructed to design creatures divergent from the examples provided to them (Smith et al., 1993, Experiment 3). When a delay intervenes between exposure and the two tasks of recalling one's own items or gener-

ating brand new items, inadvertent plagiarism increases (Brown & Halliday, 1991; Marsh & Bower, 1993). Presumably, when the problem context is reinstated, previously generated items accrue some associative strength in memory over baseline strength, thereby making them more available to be claimed as one's own contribution.

A slightly different account based on context could also account for the conformity effects that Smith et al. (1993) discovered. Rather than features or items accruing activation upon contextual reinstatement (e.g., generating additional solutions to previously encountered puzzles), the original learning context may serve as a constraint on what participants believe to constitute a valid contribution. By this account, people may abstract the features common to the examples and form a category of "space creatures" that serves as a comparative against which they judge their own novel productions as being either adequate or inadequate. If this were true, the common features remain highly "activated" by virtue of their role in comparing one's novel design against a category representation defined by those common features.

By which mechanism common features accrue activation is unclear; but, that activated information is undeniably incorporated into people's novel productions (Brown & Halliday, 1991; Brown & Murphy, 1989; Marsh & Bower, 1993; Marsh & Landau, 1995). Its inclusion, however, may not actually constrain creativity. Smith and his colleagues have argued that the heightened accessibility of an exemplar and its features either blocks (i.e., inhibits) the retrieval of other exemplars and their features or inhibits a different, more creative planned approach to the task (Smith, 1995; Smith et al., 1993). If this were true, one might expect to find that fewer creatures are generated when examples are provided than when they are not; but this was not the case across Smith et al.'s (1993) three experiments. A delicate balance clearly exists between (1) the facilitory effects of providing examples, analogies, and reminders (see, e.g., Gick & Holyoak, 1980; Ross, Ryan, & Tenpenny, 1989) and (2) the cognitive fixation (see, e.g., Smith & Blankenship, 1991) or constraining effects on creativity that are the focus of present concern. If creativity must ultimately be judged *relative* to something else as a yardstick (see, e.g., Boden, 1991; Ward, 1994), the constraining effects of examples should also be judged *in relation* to some yardstick. Defining creativity and developing measurements of it are difficult tasks. Smith et al.'s (1993) solution was to count the number of creatures produced and to assess the degree (i.e., proportion) to which common features of the examples were incorporated. These two measures may simply be insufficient metrics.

For example, in designing a novel solution to some problem, incorporation of the features provided in examples (i.e., conformity) is neither a wholly sufficient nor a satisfying demonstration that creativity has been constrained. If people in Smith et al.'s (1993) example conditions designed creatures that possessed a greater num-

ber of features other than the critical ones as compared with controls, the examples could have arguably facilitated more elaborate designs and hence *enhanced* creativity. In contrast, if the designs of the example groups were not more elaborate but nevertheless possessed more critical features of the provided examples, it becomes important to learn *of the remaining features* whether the novel or the more common features were inhibited or excluded. To our knowledge, previous research has not examined either of these possibilities. Rather, only the proportion of critical features was examined. Critically comparing the balance of features incorporated from the examples with the remaining features of people's novel creations was one goal of the experiments reported here. In this way, we had several metrics of creativity that fulfill the criterion of measuring creativity in relation to something else. Our other goals were as follows.

If features of experimenter-provided examples retain activation from presentation, greater activation of the critical features should result from seeing a greater number of examples. This prediction was tested in Experiment 1, where the number of examples was varied from 0 to 9. If greater activation forms the basis of greater retrieval blocking, conformity should rise significantly as more examples are shown. If this activation and retrieval blocking account is indeed correct (e.g., Smith et al., 1993), any given features of examples that are shown to people should be able to be "primed" for inclusion into their novel creations. This notion was tested in Experiment 2, where a pair of naturally co-occurring features was compared with a pair of artificial features that do not generally co-occur in the real world. If common features retain activation because they define a new category, conformity should be greater in the natural condition than in the artificial condition, because people should have more difficulty in learning an artificial covariation (see, e.g., Murphy & Allopenna, 1994). Another goal of Experiment 2 was to examine differences in an independent rating of subjective creativity for groups shown examples as compared with a group that was not. Finally, Experiment 3 tested whether activation of the examples would dissipate sufficiently after a 24-h delay to free people from these conformity effects. If it should do so, then an activation and blocking account of constraints on creativity would remain viable. If not, alternative theoretical accounts would need to be explored. These issues will be developed more fully in the General Discussion section.

EXPERIMENT 1

The goal of Experiment 1 was to assess whether presenting a greater number of examples that all share common critical features would increase the activation of those features. By Smith et al.'s (1993) account, seeing more examples might increase retrieval blocking and increase conformity to the examples. In a similar vein, seeing

more examples should improve or hone the categorical knowledge of what features define an "alien creature," thereby increasing conformity. This rise in conformity could be predicted by a better category representation that would result from more opportunities to abstract the critical features, and from the critical features accruing activation when one's novel creation is compared against categorical knowledge in order to determine the adequacy of one's creative contribution.

Method

Participants. Ninety University of Georgia undergraduates volunteered in return for partial course credit. Ten preliminary participants generated a pool of creature drawings that were used as the basis of the examples provided to the 80 experimental participants. Depending on when they arrived at the laboratory, 16 experimental participants were assigned to one of five between-subjects conditions varying the number of experimenter-provided examples. People were tested in small groups that ranged from 4 to 8 people.

Materials and Design. A total of nine experimenter-provided creatures was needed. Three of these were taken from Smith et al. (1993, p. 840). Six additional creatures were adapted from creations generated by the 10 preliminary participants, who received instructions identical to those in the control condition to be described shortly. As in previous work, each feature of the experimenter-provided example was labeled, and each was accompanied by a short description (examples are given in Figure 1). Following Smith et al. (1993), all nine example creatures had three attributes in common: (1) four legs, (2) two antennae, and (3) a tail. Each creature was drawn from both front and side views and was presented to people by means of an overhead projector. For the drawing task, participants were given prepared booklets containing white drawing pages separated by sheets of dark brown paper to prevent them from seeing or revisiting their earlier creations. The experimental design consisted of one between-subjects factor that specified the number of examples that people saw. That factor had five levels, specifying 0 (control), 1, 3, 6, or 9 examples.

Procedure. At the start of the experimental session, all participants were read the following introduction (adapted from Smith et al., 1993):

Imagine a planet just like Earth existing somewhere else in the universe. It is currently uninhabited. Your task is to design new creatures to inhabit the planet. Within the approximately 20 minutes allotted, draw as many new and different creatures of your own creative design as you are able. Duplication of creatures now extinct or living on the planet Earth is not permitted.

After these instructions, four groups of people studied 1, 3, 6, or 9 experimenter-provided examples for 90 sec. For all experimental groups other than the nine-example group, examples were always a subset of the next highest group (i.e., the example in the one-example group was shown in the three-example group, the examples in the three-example group were a subset of those shown in the six-example group, and the examples shown in the six-example group were a subset of the examples shown in the nine-example group). However, each person in each condition saw the identical set of examples. After the viewing period, people were admonished not to copy the examples with the following instructions, patterned after Smith et al.'s "diverge" condition (1993, Experiment 3):

The example(s) we just showed you were created by participants in another group that we tested earlier. We showed you these examples in order to help you think about creating your own original creatures and to get your creative juices flowing. However, we do not want you to copy any aspect of the examples. Please concentrate on creating your own original creations.

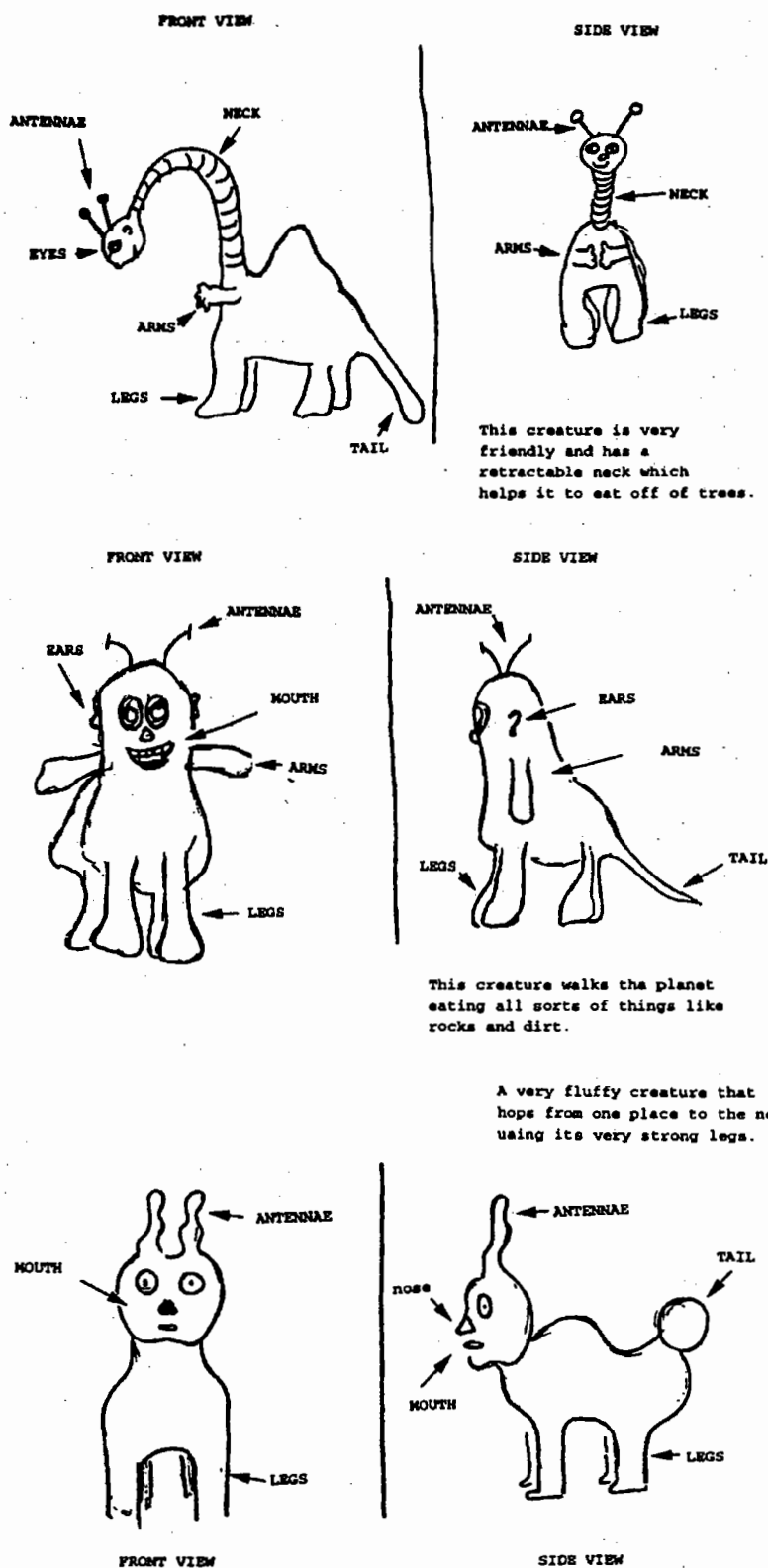


Figure 1. An example of the creatures provided to participants in Experiment 1.

Although people in the control group did not see any examples, they were asked to sit quietly for 90 sec before drawing, to equate the time given to the people in the four other conditions who did. All participants then drew creatures of their own novel design for 25 min. The booklets provided to people reminded them to provide both side and front views of each creature, to label each part of it, and to provide a one- or two-sentence description of it. After completing a design, participants continued to the next drawing page in their booklet to design a new creature. The experimenter provided additional booklets to the people who exceeded the 10 creatures that could be accommodated per booklet; but in practice, this was rarely necessary.

Results and Discussion

The creatures were scored independently by two raters who were blind to the objectives of the experiment. Each was instructed to score each creature for *every* feature it contained, regardless of whether the person had explicitly labeled it or not. After scoring, each was instructed to develop a master list of features and to rescore all of the creatures again to verify their own accuracy. There was almost no ambiguity about the features, and, on the basis of the second scoring, interrater reliability was .96 (disputes were settled in conference between the two raters and the second author). The results reported are those based on the postconference agreement of the raters and are summarized in Tables 1A–1B. When linear trends are reported, they have been adjusted for unequal spacing between the numbers of examples provided to the five groups of people.

As can be seen in the first row of Table 1A, there was a small and insignificant rise in the total number of creatures generated by the groups as the number of examples shown to them increased [$F(4,75) < 1$]. The second row of that table displays a similar pattern in the average number of features that constituted people's creations [$F(4,75) = 1.40, p > .10$]. These two rows of data, therefore, indicate that creativity as measured by either total output or elaborateness of design did not vary with the number of examples that people saw. In contrast, the first five rows of Table 1B indicate that as the number of examples increased, so did people's propensity to incorporate the three critical features of antennae, four legs, and a tail. The first row contains the overall conformity score used by Smith et al. (1993) and was calculated for a participant as the average proportion of the three critical features contained in their creations. As predicted, conformity rose significantly as the number of examples increased [omnibus $F(4,75) = 5.91, p < .01$; linear trend $F(1,75) = 4.75, p < .01$]. Compared with Smith et al.'s participants (e.g., 1993, Experiment 1), these people produced on average one additional creature (because they had 5 min of additional drawing time), but conformity in the control and three-example conditions was remarkably similar to the findings in that earlier work. Rows 2–4 display individual conformity scores by feature (e.g., the proportion of creatures possessing a tail, etc.) and were averaged to arrive at row 1, the overall conformity score. As expected, con-

Table 1A
Total Number of Creatures and Mean Number of Features
(With Standard Deviations) in Experiment 1

	No. of Experimenter-Provided Examples									
	0		1		3		6		9	
	No.	SD	No.	SD	No.	SD	No.	SD	No.	SD
Total creatures	4.4	1.8	5.1	1.7	4.9	1.0	5.2	2.2	5.4	2.5
Mean features	6.6	1.6	6.3	2.5	6.9	1.6	7.8	2.2	7.1	1.7

Table 1B
Average Conformity and Proportion of Features by Type
(With Standard Deviations) in Experiment 1

	No. of Experimenter-Provided Examples									
	0		1		3		6		9	
	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD
Conformity	.087	.083	.137	.103	.189	.126	.233	.141	.280	.161
Antennae	.072	.117	.120	.177	.213	.231	.149	.228	.206	.212
Tail	.122	.159	.202	.187	.242	.194	.353	.270	.405	.338
Four legs	.066	.111	.088	.128	.111	.125	.199	.195	.228	.259
Critical Features	.044	.047	.068	.069	.076	.048	.093	.046	.121	.067
Noncritical Features	.956	.047	.932	.069	.924	.048	.907	.046	.880	.067
Novel	.056	.082	.045	.059	.070	.091	.044	.057	.053	.054
Common	.807	.091	.772	.158	.752	.125	.724	.099	.674	.125
Uncommon	.083	.054	.098	.110	.094	.069	.135	.065	.147	.105
Miscellaneous	.010	.016	.016	.043	.007	.016	.004	.015	.006	.015
Examples										
Novel			.000		.000		.000		.014	
Common			.625		.625		.651		.622	
Uncommon			.000		.000		.000		.022	
Miscellaneous			.000		.000		.000		.000	

formity generally rises linearly in all three cases [omnibus for tail $F(4,75) = 3.65, p < .01$, linear trend $F(1,75) = 3.75, p < .01$; omnibus for four legs $F(4,75) = 2.73, p < .05$, linear trend $F(1,75) = 3.25, p < .01$; omnibus for antennae $F(4,75) = 1.46, p = .224$, linear trend $F(1,75) = 1.70, p = .09$].

The fifth row of Table 1B sets forth the average proportion of critical features that constituted people's creations. For example, the control group averaged 6.6 features per creation, 4.4% of which were critical features; the nine-example group averaged 7.1 features per creation, 12.1% of which were critical features, and so forth. As is evident, the proportion of critical features that people incorporated increased with the number of examples they saw [omnibus $F(4,75) = 4.15, p < .01$; linear trend $F(1,75) = 3.97, p < .01$]. Because the composition of the "non-conforming" features of people's novel creations is of particular interest, the sixth row of Table 1B sets forth the average proportion of features that were noncritical (i.e., rows 5 and 6 sum to 100% for each column). Because the proportion of features that were critical (row 5) and those that were not (row 6) are linearly dependent, statistical analysis is redundant. In the next four rows, however, the noncritical features were classified into four categories: (1) *novel* features that are generally not found on animals (see Ward's definition, 1994, pp. 7–8), (2) *common* features that are found on many species of animals, (3) *uncommon* features that are generally restricted to a small class of animals, and (4) *miscellaneous* features that are largely confined to articles of clothing, earrings, and so forth. Because this classification is critical to interpretation of the results, it has been included in the Appendix.

Unlike conformity, the number of novel features that were incorporated into people's designs did not vary with the number of examples they saw [$F(4,75) < 1$]. In contrast, the number of features commonly found on Earth animals declined significantly as the number of examples increased [omnibus $F(4,75) = 2.73, p < .05$; linear trend $F(1,75) = -3.25, p < .01$]. Although increasing the number of examples slightly increased the number of uncommon animal features, the rise was small and unreliable [$F(4,75) = 1.80, p > .13$]. In general, the miscellaneous/clothing category contributed little to people's designs and neither varied with the number of examples presented [$F(4,75) < 1$] nor changed the pattern of results when pooled with the novel features. Because people seem to define a novel creature as constituting a constant 6 to 7 features (Table 1A), as people increasingly incorporated the critical features of the examples they were shown (Table 1B), rather than at the expense of the novel or uncommon features, fairly common attributes of animals were the ones that experienced a decline (Table 1B). Together these results suggest that the provision of examples may not actually constrain creativity (as measured by the proportion of novel and uncommon features included in people's drawings), but rather, influence which of many more common attributes people choose to include or to exclude.¹

One concern is whether or not the examples might have differed in any systematic way across the four groups who saw them. The last four rows of Table 1B set forth the proportion of the total number of features in our examples that correspond to the four categories of noncritical features. Visual inspection of the table suggests neither a decline in the proportion of common features in our examples nor any other uniform trends that might account for the results just described.

EXPERIMENT 2

The results from Experiment 1 suggest that, as more examples are presented to people, critical features of those examples have an increased probability of being incorporated into their novel designs. At face value, this conformity to experimenter-provided examples would suggest a constraint to creativity. Because people who saw an increasingly greater number of examples neither generated more designs nor generated more elaborate designs, the constraint seems to be isolated to their propensity to exclude more common, everyday features of animals—and not to exclude more novel and uncommon features.

One purpose of the next experiment was to generalize the results of Experiment 1. Because a tail and four legs tend to covary in Earth animals (see below), that combination could largely inflate the conformity effect; that is, placing four legs on an animal may increase the probability of also placing a tail on it. The goal of Experiment 2 was to ascertain whether introducing an artificial covariation in the examples would lead to conformity in people's novel creations. A simple activation account followed by retrieval blocking predicts that any feature should be able to accrue activation and inhibit more creative solution paths. However, if critical features are more available because they define a category against which participants feel compelled to compare their novel designs, then conformity may be larger in the natural condition than in the artificial. This prediction is based on the ease with which categories are learned. People generally have greater difficulty learning categories in which the covariation of features conflicts with real-world knowledge (see, e.g., Murphy & Allopenna, 1994). The examples in Experiment 2 covaried either antennae and claws (artificial) or four legs and tail (natural). The simple prediction is that people in the artificial condition should incorporate antennae and claws with a greater probability than those in the natural or control conditions; and people in the natural condition should tend to incorporate four legs and tail more often than those in the artificial condition or a control group. In other words, conformity to specific examples should be greater in the experimental conditions than in a control condition. An independent assessment of subjective creativity was performed to test the prediction that a control condition who did not see examples would be rated as more creative than either of the groups who did see experimenter-provided examples. As

in Experiment 1, a final goal of these manipulations was to examine the noncritical features of people's drawings under these three conditions.

Method

Participants. Ninety-five University of Georgia undergraduates volunteered in exchange for partial fulfillment of a course requirement. In a manner similar to that in Experiment 1, 25 people were randomly assigned to each of three between-subjects conditions on the basis of their arrival at the laboratory. These participants were tested in small groups that ranged from 6 to 12 people, and none had participated in Experiment 1. The additional 20 participants were used in two tasks involving ratings of subjective creativity as described in the Results and Discussion section.

Materials and Design. A total of 12 examples were created by choosing and adapting some of the designs generated by people in Experiment 1. Six examples contained two features that generally are negatively correlated in nature: two antennae and claws.² The remaining six examples contained two features that generally do positively covary in nature: four legs and a tail. Evidence for this natural covariation comes from the fact that 98% of the first 42 animals listed in Battig and Montague's (1969) norms for four-footed animals possess a tail. As in Experiment 1, the examples that were created listed the features of the creatures and provided a brief description of them.

Participants in the *artificial* condition saw only the six creatures that possessed antennae and claws, whereas people in the *natural* condition studied the six creatures in which four legs and a tail were covaried. Each example was drawn to show front and side views. Participants drew their creations in booklets identical to those used in Experiment 1. A control condition did not see any examples, thereby making the experimental design a single between-subjects factor designating the type of prior exposure to examples (control, artificial covariation, natural covariation).

Procedure. The procedure followed that of Experiment 1 almost identically. The same introductory remarks were given to all people. People in the artificial and natural conditions then saw their respective six examples presented for 30 sec each (3 min total). All participants were admonished (using the previous instructions) not to copy the examples that they saw. People in the control condition did not see any examples and were asked to sit quietly for 3 min while awaiting further instructions. All participants were given 25 min to draw as many novel creatures as they could.

Results and Discussion

Creatures were scored in the same manner as in Experiment 1 by the same two raters (still blind, at this point, to the objectives of the experiments). Interrater reliability was .98, slightly higher than in Experiment 1, probably owing to their experience of twice scoring the creatures generated in that prior experiment. The results are summarized in Tables 2A–2B. Where specific statistical contrasts support a claim better, they are reported in addition to omnibus tests.

As in Experiment 1, neither the total number of creatures generated by the two experimental conditions [row 1 of Table 2A; $F(2,74) < 1$] nor the complexity of those designs [row 2; $F(2,74) = 1.66, p > .10$] differed from performance for the control condition. In contrast to that equivalence of the three conditions, the conformity data in Table 2B suggest that different patterns of conformity were displayed by the groups. Conformity was scored in two ways: (1) a *strict* scoring measured the proportion of creatures that possessed *both* critical features (Table 2B,

Table 2A
Total Number of Creatures and Mean Number of Features
(With Standard Deviations) in Experiment 2

	Testing Condition					
	Control		Artificial		Natural	
	No.	SD	No.	SD	No.	SD
Total creatures	4.84	2.39	4.48	1.19	5.04	2.01
Mean features	5.87	1.68	6.70	1.73	6.58	1.77

Table 2B
Average Conformity and Proportion of Features by Type
(With Standard Deviations) in Experiment 2

	Testing Condition					
	Control		Artificial		Natural	
	Avg.	SD	Avg.	SD	Avg.	SD
Conformity						
Strict						
Tail/legs	.033	.070	.044	.094	.129	.163
Antennae/claws	.000	.000	.036	.085	.000	.000
Lenient						
Tail/legs	.193	.215	.257	.229	.356	.270
Antennae/claws	.134	.194	.254	.246	.139	.198
Critical Features						
Tail/legs	.030	.034	.037	.037	.070	.056
Antennae/claws	.022	.034	.039	.040	.017	.025
Noncritical Features						
Novel	.095	.102	.068	.087	.064	.082
Common	.685	.136	.699	.110	.669	.150
Uncommon	.150	.103	.131	.076	.146	.109
Miscellaneous	.014	.032	.010	.024	.026	.068

rows 1 and 2), and (2) a *lenient* scoring measured the proportion displaying *either* critical feature (i.e., the conformity score reported in Experiment 1 and set forth in rows 3 and 4 of Table 2B). By a strict scoring, people in the natural condition produced more creatures that had tails and legs than did those in the artificial and control conditions [omnibus $F(2,74) = 5.18, p < .01$; contrast $t(72) = 3.2, p < .01$] and only people in the artificial condition produced creatures that had both antennae and claws [omnibus $F(2,74) = 4.50, p < .01$; contrast $t(72) = 3.00, p < .01$]. These results suggest that a covariation of two features, whether natural or artificial, is adopted from the examples into participants' novel designs. A lenient scoring weakened these effects, but creations by people in the natural condition nevertheless tended to contain four legs or a tail more often than those produced by either the artificial or control conditions [$F(2,74) = 2.95, p = .06$; contrast $t(72) = 2.24, p < .05$], while antennae or claws were included more often by people in the artificial condition than by those in the other two groups [omnibus $F(2,74) = 2.51, p = .08$; contrast $t(72) = 2.24, p < .05$].

As in Experiment 1, the average proportion of features that were either *critical* (i.e., shown in the examples) or *noncritical* was analyzed (see Table 2B). The proportion of total features that were four legs and tail was higher in the natural condition than in the other two groups [row 5; omnibus $F(2,74) = 6.01, p < .01$; contrast $t(72) = 2.84, p < .01$]; and there was a trend for antennae and claws to

constitute a greater proportion of features in the artificial condition than in the other groups [row 6; omnibus $F(2,74) = 2.81, p = .07$; contrast $t(72) = 2.75, p < .01$]. Of the noncritical features, there was no trend for any of the novel, common, uncommon, or miscellaneous (rows 7-10, respectively) to differ among the three groups [all $F_s(2,74) < 1$]. However, close inspection of those rows suggests that the two experimental conditions used fewer novel and uncommon features than did the control condition, but not reliably so. As such, the conclusion can be drawn that no one category of the noncritical features contributed disproportionately to the constraints observed in the experimental groups as compared with the controls.

In terms of disambiguating a retrieval blocking account from an account in which critical features remain available because people compare their novel designs against a set of abstracted features of the examples, conformity was arguably greater in the natural condition (12.9% and 35.6% by strict and lenient scoring, respectively) than in the artificial condition (3.6% and 25.4%). There is ample evidence from the categorization literature (e.g., Murphy & Allopenna, 1994; Murphy & Wisniewski, 1989) to demonstrate that people have difficulty in learning categories in which the features conflict with their real-world knowledge. For example, people had trouble learning the features "has a sharp point" and "used for smoothing wallpaper" for a category as compared with one in which the second attribute was "used for stabbing bugs." The fact that people tended to incorporate more features in the natural condition than in the artificial condition might suggest that they are forming a category "creatures to inhabit a distant planet" when shown the experimenter-provided examples. We will return to these issues in greater detail in the General Discussion section.

Additional analyses were conducted to test whether providing examples resulted in less creative output as subjectively rated by people blind to the experimental aims. Four people were recruited to rate on a 5-point Likert scale each of the creations produced across all the conditions. If providing examples inhibits creativity, the control condition might be expected to have the highest mean creativity rating. In fact, the artificial condition had the highest mean creativity rating (3.51, $SD = .64$), followed by the natural group (3.07, $SD = .48$) and the control condition (3.08, $SD = .68$). The differences between the artificial group and the other two conditions was significant [$F(2,72) = 4.15, p < .05$], suggesting that providing examples can have a different sort of influence on creative production. Perhaps the artificial condition was rated as more creative because these participants copied the rather bizarre covariation of antennae and claws; and perhaps the examples shown to the natural condition did not provide additional ideas beyond what control participants were able to devise on their own without any guidance. To ascertain whether our artificial examples were subjectively rated as more creative than our natural examples, 16 people were recruited to rate all 12 examples on a 5-point Likert scale (presented to them in a random order). Interestingly, the mean cre-

ativity rating for the artificial examples (3.29, $SEM = 0.16$) was very close to that for the natural examples (3.21, $SEM = 0.24$) that we showed participants [$t(15) < 1$]. Together, these analyses suggest a genuine influence of providing examples on people's creative output, and that influence is not due to the examples' being subjectively more creative in the artificial condition.

EXPERIMENT 3

Together, Experiments 1 and 2 suggest that features of examples, whether correlated in nature or not, will be incorporated into people's novel creations. The results of Experiment 2, however, were more consistent with an activation (or categorization) account than a retrieval blocking account. If retrieval blocking was the only operative mechanism, conformity should have been equivalent in the natural and artificial conditions, which it was not. The purpose of Experiment 3 was to again explore the hypothesis offered by Smith et al. (1993) that retrieval blocking might constitute the locus of the constraints observed in the present experiments and those conducted before (i.e., Smith et al., 1993). Smith and his colleagues have repeatedly shown that fixation effects due to heightened activation of competing material can be greatly reduced by the introduction of a delay before assessing performance (an excellent review is contained in Smith, 1995). In fact, Smith et al. (1993, Experiment 2) introduced a 23-min delay after exposure to three experimenter-provided creatures before drawing; but the manipulation had no consequent effect on conformity. Memory for pictures, however, is quite good even after fairly long delays (Shepard, 1967), which may have mitigated any effect of a short delay in this paradigm.

Therefore, the method of Experiment 3 was simply to use a 1-day delay, after which any heightened activation of the examples should have subsided. On the one hand, if a retrieval blocking account is correct, less conformity should be observed at delayed test than at immediate test (just as Smith et al., 1993, had predicted). On the other hand, if the conformity observed in such studies is conceptually similar to the unconscious plagiarism people commonly exhibit with verbal materials (e.g., Brown & Halliday, 1991; Marsh & Bower, 1993; Marsh & Landau, 1995), a delay should increase conformity for all of the reasons cited in our introductory remarks. Finally, a categorization account in which people abstract the common features of the examples and use them as the basis of a category—say, "space creatures"—makes slightly different predictions with respect to a delay. If the newly learned category representation is weakened or lost over the course of a delay, less conformity should be observed. If it is not, equivalent conformity should be observed at both immediate and delayed testing.

Method

Participants. Seventy-four University of Georgia students were recruited from the same pool as before. None had served in the pre-

vious experiments. Thirty-six were tested in an *immediate* condition, and 38 were tested in a *delay* condition, as detailed shortly.

Materials and Design. The procedure was similar to that of the previous experiments. Three experimenter-provided examples were chosen that all contained the three critical features of four legs, a tail, and antennae. These were shown to people on an overhead projector for 90 sec.

Procedure. After viewing the examples, people in the delay condition were dismissed and required to return the following day. Participants in both conditions were given identical instructions (see Experiment 1) admonishing them not to copy the examples but to draw creative creatures of their own design. Both groups drew for 25 min, the principle difference between them being that only the instructions intervened between viewing the examples and drawing for the immediate group, whereas a full day and the instructions intervened for the delay group.

Results and Discussion

The critical results are set forth in Tables 3A–3B. Consistent with the three-example condition in Experiment 1 and the natural condition of Experiment 2, about five drawings were made in 25 min and about six features constituted each creation. Neither the number of creatures [$F(1,73) < 1$] nor the mean number of features [$F(1,73) < 1$] differed between immediate and delayed testing (Table 3A). Smith et al.'s (1993) conformity score, however, does differ (Table 3B); conformity appears to rise after a 1-day delay [$F(1,73) = 4.74, p < .05$]. Similarly, using our metric of taking the proportion of total features that were the critical features of the provided examples; there was an increase from immediate to delayed testing [$F(1,73) = 5.74, p < .05$]. As in the previous experiments, the remaining (i.e., noncritical) features were examined in order to isolate exactly which types of features tended to be excluded at delayed testing. As can be

seen in Table 3B, small, insignificant declines were evident in the novel, uncommon, and miscellaneous categories [all $Fs(1,73) < 2.5, ps > .10$]. Thus, the increase in conformity to experimenter-provided examples that follows a delay does not appear to result from excluding any particular class of noncritical features. How the results of this experiment, and those of the previous two, constrain the various theoretical accounts will now be considered in some detail.

GENERAL DISCUSSION

The foregoing results support the claim that, in creative activities, providing examples may ultimately alter the nature of the creative product (see, e.g., Smith, 1995; Smith et al., 1993; Ward, 1994, 1995). Because aspects of experimenter-provided examples were included in novel designs at rates higher than those in control conditions (Experiments 1 and 2), the results also support Smith et al.'s claim that creativity has been "constrained." Across tasks and conditions, however, the manipulations in Experiments 1–3 resulted in neither a different total number of designs created (i.e., a measure of productivity) nor a different mean number of features (i.e., one measure of complexity). The lack of any differences afforded a close examination of the noncritical features; and, those analyses did not generally show that novel or uncommon features were excluded at the expense of including "primed" features contained in the examples. Rather, there were general declines among the four types of noncritical features (Experiments 2 and 3) or a significant decline in common features (Experiment 1). Moreover, the subjective, yet independent, rating of creativity in Experiment 2 showed that "primed" experimental participants in the artificial condition were viewed as more creative than control participants. At face value, these results from using different metrics of creativity contravene the conclusion that providing examples constrains creativity.

Before we consider the theoretical explanation of these results, a potential criticism of these experiments must be addressed—namely, that the admonition instruction given to each of the examples groups not to copy may have served as a reminder or hint to reflect back on the examples, thereby increasing conformity. Although this may be possible, specifically manipulating different sorts of instructions to participants usually has little effect in this paradigm. For example, Smith et al. (1993, Experiment 3) compared four conditions, including a group specifically asked to *diverge* from the examples that they saw, a group told to *converge* (i.e., copy), a group given neither converge nor diverge instructions, and a control condition that saw no examples. Conformity occurred in all groups who saw examples, but diverge instructions did not decrease conformity in comparison with instructions that made absolutely no reference to the examples (i.e., the standard group). As mentioned earlier, Ward (1994, Experiment 3) has also manipulated instructions by asking people to draw creatures "beyond their wildest imagination" with no consequent effect on creative output as

Table 3A
Total Number of Creatures and Mean Number of Features
(With Standard Deviations) in Experiment 3

	Testing Condition			
	Immediate		Delay	
	No.	SD	No.	SD
Total creatures	5.47	2.20	5.03	1.78
Mean features	6.25	1.95	6.52	1.46

Table 3B
Average Conformity and Proportion of Features by Type
(With Standard Deviations) in Experiment 3

	Testing Condition			
	Immediate		Delay	
	Avg.	SD	Avg.	SD
Conformity	.169	.13	.242	.16
Antennae	.115	.14	.262	.29
Tail	.235	.21	.270	.28
Four legs	.158	.19	.195	.24
Critical Features	.078	.06	.115	.07
Noncritical Features	.922	.06	.885	.07
Novel	.082	.07	.066	.07
Common	.705	.11	.716	.12
Uncommon	.119	.10	.089	.06
Miscellaneous	.012	.03	.007	.03

compared with conditions not given that instruction. That said, merely providing people with examples may affect how they go about the drawing task. Given the problem-solving task of designing a creature, people who saw examples may remind themselves about the examples that they saw and thereby tend to incorporate the critical features. This account is consistent with demonstrations by Ross (1987) that the more similar some earlier learned information (i.e., the examples that people saw) is to a current problem (i.e., the drawing task), the greater the probability that people will access (and sometimes use) that earlier learned information.

From a theoretical perspective, aspects of presented examples may be incorporated into novel designs because they have retained activation from their earlier presentation. However, that the heightened activation causes retrieval blocking (Smith et al., 1993) or less creative solution paths to be explored (Smith, 1995) is somewhat less clear. In Experiment 2, more conformity was found in the artificial than in the natural condition. A retrieval blocking account cannot explain that effect; retrieval blocking would predict equal conformity. In addition, activation should have dissipated after the 24-h delay introduced in Experiment 3, and that should have alleviated some of the blocking. But, contrary to that prediction, conformity was significantly *greater* in the delay condition than in the immediate one. That rise in conformity is predicted by models and theories of inadvertent plagiarism (e.g., Brown & Murphy, 1989; Marsh & Bower, 1993; Marsh & Landau, 1995). Prior investigations of cryptomnesia are similar to the paradigm used here because people are first exposed to material and then admonished not to use it in their creative problem-solving tasks. In that sense, then, these studies document the inadvertent theft of pictorial materials. To explain the increase in cryptomnesia, the model proposed by Marsh and Bower (1993) argues that when the drawing context is reinstated after a day's delay, aspects of the examples are activated above baseline concepts in memory. That activation could be sufficient for certain aspects to be brought to mind, but insufficient to warrant rejection as "part of an example I am not supposed to copy."

If this activation account were viable, then determining at what level of abstraction the activation accrues is critically important. That is, is the accrual of activation at the level of memory for an entire presented example or merely at the level of features of those examples? Experiments 1–3 report cryptomnesia at the level of features, whereas earlier investigations demonstrated cryptomnesia only for entire items. In Ward's (1994, 1995) structured imagination theory, people bring to mind an exemplar and then attempt to modify it in novel ways. Does the influence of past experience on drawing novel creatures come from bringing to mind an entire presented example, or does it arise in the "modification stage" of that theory? As yet, the critical experiments have not been conducted in which people attempt to recall the examples that they were shown as a means of demonstrating what they have retained from that prior exposure.

However, there is an interesting distinction, in that Ward's theory accounts for plagiarism by explicit, deliberate recollection of an exemplar, whereas in models of unconscious plagiarism, the influence of prior information is assumed to be implicit by virtue of the admonition given to people not to copy what they have been exposed to earlier.

An activation account of the present results also raises several other unresolved issues. For example, if the heightened activation of critical features blocks retrieval in some way, does it block the retrieval of a class of "animal features," does it block other solution paths, or does it block both? These questions are difficult to answer, because in most studies of retrieval blocking, specific "correct" items on a previously learned list could be assessed as having been recalled or not (e.g., Roediger, 1974). Similarly, in investigations of retrieval blocking in problem-solving tasks (e.g., Remote Associates, TOT solutions), the problems have uniformly had one correct solution (e.g., Smith, Balfour, & Brown, 1994; Smith & Blankenship, 1989, 1991). In contrast, people who participated in these experiments were free to use any of at least 65 features (the number on the master coding list) and assemble them in almost any fashion they chose; as such, there was no single correct solution. Moreover, had they used features in their novel creations from common household objects (e.g., faucets, brooms, string, etc.), the list of possible features would become limitless. Just as knowing what something is not is a poor retrieval cue, knowing what features one's innovation should not include is a poor indicant of what that novel design should include.

In contrast to the heightened activation interpretation offered by Smith et al. (1993) and Marsh and Bower's (1993) model of inadvertent plagiarism, a promising account of these experiments is a variant of Ward's theory of structured imagination. By this new account, people abstract the common features of the examples that they see in order to arrive at a new category, such as "creatures that inhabit a distant planet." Because people are rather sensitive to correlated category attributes (e.g., Rosch, 1978) and are able to reliably detect fairly small correlations in the laboratory (e.g., Heit, 1992; Lewicki, Hill, & Czyzewska, 1992), the provision of examples may serve to define (or redefine) the category and, therefore, what constitutes an acceptable new exemplar of that category. To state this slightly differently: the examples may define the context in which permissible exemplars are generated. This context, once defined, can be reinstated (even a day later, as in Experiment 3) and affect creative output. This notion is consistent with models of category learning that suggest that better learning should be associated with exposure to an increasing number of exemplars (e.g., Clapper & Bower, 1994; Nosofsky, 1988).

If providing more examples increasingly hones knowledge of category structure, it may also increasingly constrain what constitutes an acceptable novel addition to that category, thereby constraining creativity; and this is just the result that was observed in Experiment 1. Further credence is lent to such an explanation by the fact that

people tend to "define" a novel creature as having a stable number of features (about five or six). The foregoing analysis complements Heit's (1994) work on category learning. Heit suggests that when people observe exemplars of new categories, they also retrieve prior examples from previously known categories (for a related argument, see Spalding & Ross, 1994). Although these prior examples that are brought to mind do guide or constrain new category learning, they do not prevent it. These notions of contextual or categorical definitions are currently being explored and refined. For example, an activation account of the results reported in the present experiments can partially be disambiguated from a category definition account by changing the modalities at input and test. This sort of demonstration relies on the fact that activation often does not cross modalities in tests of implicit memory (see, e.g., Bassili, Smith, & MacLeod, 1989). Another demonstration would be to compare groups who receive the same features (e.g., legs, tails, pointed ears) but place them on "space creatures" for one group and in the context of Earth animals for a second group. Presumably, if the examples serve to define a new category, people who are shown Earth animals (for which they already have preexisting categories) should show less conformity.

The preceding analysis notwithstanding, the results of Experiment 3 do not match the predictions of such a category redefinition hypothesis. Presumably, the fate of a newly learned category after a 1-day delay is either that it remains stable and intact, or that it is weakened in memory when the context is reinstated after the delay. By this account, conformity (i.e., plagiarism) should either remain stable or decrease after a delay. But, it did neither; it rose significantly. The only reconciliation of the results from Experiment 3 and a categorization account would be to assume that critical features are by some means sharpened or focused in memory after a delay as compared with noncritical ones. As an alternative account to any theoretical explanation involving activation, the *fluency* approach has been advocated by Jacoby and his colleagues (e.g., Jacoby et al., 1989). By this account, "memory" is an attribution that is made by assessing the independent contributions of automatic (i.e., implicit) and conscious (i.e., explicit) processes. In a manipulation similar to the delay introduced in Experiment 3, Jacoby et al. showed people lists of famous and nonfamous names. When tested after a day's delay, people tended to call nonfamous names "famous," in comparison with control groups who were tested immediately. Presumably, after a day's delay, there was less explicit memory (i.e., recollective processes) to oppose more automatic processes and that lead to more misattributions of familiarity. Activated information is processed more fluently, and more fluently processed information tends to be attributed to oneself (Marsh & Landau, 1995). When conscious recollection cannot oppose that attribution, people may assume that features that come to mind easily are a product of their own creative processes, rather

than recently experienced. Such a hypothesis is consistent with our introductory remarks that people may fail to spontaneously monitor the source of their ideas in a creative problem-solving task.

As reviewed here, the empirical and theoretical findings proposed by Smith, Ward, and their colleagues provide important insights into the cognitive processes that support creative endeavors. Once viewed entirely as an individual difference, creativity is now known to be supported by very regular, universal, cognitive processes (Smith, 1995; Smith et al., 1993; Ward, 1994, 1995). The present results provide further evidence that past experience is brought to bear in creative problem-solving tasks. Those results also suggest that the effect of memory for the past experiences may not limit or constrain creativity in any pejorative way. Of course, when a novel design incorporates a flaw or suboptimal component from past experience (see, e.g., Jansson & Smith, 1991), there may very well be a "constraint" on creativity. Similarly, when cues and examples lead people down incorrect solution paths, creativity is indeed hampered (e.g., Smith, 1995). In the present studies, however, if there is any constraint to creativity, it could result from people abstracting critical features from examples that are provided and testing their novel designs against whether or not that design is an adequate "category member." Or it could arise by means of insufficiently monitoring the original source of the features that one is incorporating into a novel design. As the study of creativity advances, which it seems to be doing at a rapid pace while avoiding the pitfalls of earlier psychological inquiry into the topic, the relative contributions and the viability of such factors as activation, retrieval blocking, and categorization will most assuredly sort themselves out.

REFERENCES

- BASSILI, J. N., SMITH, M. C., & MACLEOD, C. M. (1989). Auditory and visual word-stem completion: Separating data-driven and conceptually driven processes. *Quarterly Journal of Experimental Psychology*, 41A, 439-453.
- BATTIG, W. F., & MONTAGUE, W. E. (1969). Category norms of verbal items in 56 categories: A replication and extension of the Connecticut Category Norms. *Journal of Experimental Psychology*, 80(3, Pt. 2), 1-46.
- BODEN, M. (1991). *The creative mind: Myths and mechanisms*. New York: Basic Books.
- BROWN, A. S., & HALLIDAY, H. E. (1991). Cryptomnesia and source memory difficulties. *American Journal of Psychology*, 104, 475-490.
- BROWN, A. S., & MURPHY, D. R. (1989). Cryptomnesia: Delineating inadvertent plagiarism. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15, 432-442.
- CLAPPER, J. P., & BOWER, G. H. (1994). Category invention in unsupervised learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 20, 443-460.
- FINKE, R. A., WARD, T. B., & SMITH, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.
- GICK, M. L., & HOLYOAK, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, 12, 306-355.
- HEIT, E. (1992). Categorization using chains of examples. *Cognitive Psychology*, 24, 341-380.
- HEIT, E. (1994). Models of the effects of prior knowledge on category

- learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 20, 1264-1282.
- JACOBY, L. L., KELLEY, C. M., BROWN, J., & JASECHKO, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influence of the past. *Journal of Personality & Social Psychology*, 56, 326-338.
- JANSSON, D. G., & SMITH, S. M. (1991). Design fixation. *Design Studies*, 12, 3-11.
- LEWICKI, P., HILL, T., & CZYZEWSKA, M. (1992). Nonconscious acquisition of information. *American Psychologist*, 47, 796-801.
- MARSH, R. L., & BOWER, G. H. (1993). Eliciting cryptomnesia: Unconscious plagiarism in a puzzle task. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 19, 673-688.
- MARSH, R. L., & LANDAU, J. D. (1995). Availability in cryptomnesia: Assessing its role in two paradigms of unconscious plagiarism. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21, 1568-1582.
- MEDNICK, S. A. (1962). The associative basis of the creative process. *Psychological Review*, 69, 220-232.
- MURPHY, G. L., & ALLOPENNA, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 20, 904-919.
- MURPHY, G. L., & WISNIEWSKI, E. J. (1989). Feature correlations in conceptual representations. In I. G. Tiberghien (Ed.), *Advances in cognitive science: Vol. 2. Theory and applications* (pp. 23-45). Chichester, UK: Ellis Horwood.
- NOSOFSKY, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 14, 54-65.
- ROEDIGER, H. L., III (1974). Inhibiting effects of recall. *Memory & Cognition*, 2, 261-269.
- ROSCH, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 28-36). Hillsdale, NJ: Erlbaum.
- ROSS, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 13, 629-639.
- ROSS, B. H., RYAN, W. J., & TENPENNY, P. L. (1989). The access of relevant information for solving problems. *Memory & Cognition*, 17, 639-651.
- SHEPARD, R. N. (1967). Recognition memory for words, sentences and pictures. *Journal of Verbal Learning & Verbal Behavior*, 6, 156-163.
- SMITH, S. M. (1995). Fixation, incubation, and insight in memory and creative thinking. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 135-156). Cambridge, MA: MIT Press.
- SMITH, S. M., BALFOUR, S. P., & BROWN, J. M. (1994). Effects of practice on TOT states. *Memory*, 2, 31-49.
- SMITH, S. M., & BLANKENSHIP, S. E. (1989). Incubation effects. *Bulletin of the Psychonomic Society*, 27, 311-314.
- SMITH, S. M., & BLANKENSHIP, S. E. (1991). Incubation and the persistence of fixation in problem solving. *American Journal of Psychology*, 104, 61-87.
- SMITH, S. M., WARD, T. B., & SCHUMACHER, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition*, 21, 837-845.
- SPALDING, T. L., & ROSS, B. H. (1994). Comparison-based learning: Effects of comparing instances during category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 20, 1251-1263.
- TORRENCE, E. P. (1974). *The Torrence tests of creative thinking: Norms-Technical manual*. Bensenville, IL: Scholastic Testing Service.
- TVERSKY, B., & HEMENWAY, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, 113, 169-193.
- WARD, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, 27, 1-40.
- WARD, T. B. (1995). What's old about new ideas? In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 157-178). Cambridge, MA: MIT Press.

NOTES

1. An alternative explanation for the decline in common features is that as people tend to include the critical attribute of four legs, it precludes them from using three of the common attributes (arms, hands, and two legs). This preclusion is illusory, however, because a number of our participants' drawings contained creatures that had both four legs and arms. Nevertheless, to test this relationship, we excluded these three attributes and found that, while attenuated, the proportion of common attributes maintained its decline as more examples were shown [linear trend $F(1,75) = 2.11, p < .05$].

2. Our intuition is that these two features do not naturally covary, because the only animal we could think of that had both antennae and claws was a lobster.

APPENDIX

Classification of Features Used in Experiment 1

Novel	
Blades/knives	Propeller
Bubbles	Robotic in nature
Eyestalks	Speakers
Foliage	Springs
Light (photons)	Suction cups
Odd locomotion	Wheels
Common	
Anus	Head/neck
Arms	Internal organs (e.g., heart)
Breasts	Mouth
Claws	Nose
Ears	Paws/hooves
Eyes	Teeth
Fangs	Tongue
Feet/toes	Two legs
Fur/hair	Whiskers
Hands/fingers	
Uncommon	
Belly button	Shell
Beak	Snakelike/reptilian
Feathers	Spikes/quills/spines
Fins	Spots/dots
Flippers	Stinger
Gills	Stripes
Horns	Tentacles
Mane	Trunk/snout
3, 5, or more legs	Tusks
Pouch	Webbed feet
Scales	Wings
Segmented body	
Miscellaneous	
Belt	Hat
Cape	Shoes
Earrings	Tie

(Manuscript received May 31, 1995;
revision accepted for publication August 26, 1995.)