

# The inadvertent use of prior knowledge in a generative cognitive task

RICHARD L. MARSH

*University of Georgia, Athens, Georgia*

THOMAS B. WARD

*Texas A&M University, College Station, Texas*

and

JOSHUA D. LANDAU

*University of Georgia, Athens, Georgia*

In four experiments with 332 participants, participants were asked to generate novel nonwords for English categories. When participants were shown examples embedded with regular orthographic structures, participants' nonwords tended to conform orthographically to the examples, despite instructions to avoid using features of the examples. The effect was found with immediate testing (Experiment 1) and delayed testing (Experiment 2). The effect was also found with arbitrary features (Experiments 1-4), as well as with naturally occurring orthographic regularities (Experiment 4). Participants had difficulty avoiding the use of this prior knowledge, despite being able to list the features they were asked to avoid (Experiment 3). The results are discussed in terms of the inadvertent use of prior knowledge in generative cognitive tasks.

Broadly defined, creative cognitive tasks are ones that require participants to design novel products. Such tasks are commonplace and include a range of ordinary human activities, such as devising new plans, new consumer goods, new melodies, new works of literature, and a wide variety of other scientifically and practically useful contributions. Despite the ubiquity of these generative endeavors, however, surprisingly little research has focused on the basic cognitive processes involved in formulating novel constructions. Our interest in studying creative cognition stems from our desire to understand those basic processes.

Although there certainly must be differences across domains in the way participants reason when they create novel entities (see, e.g., Greeno & Simon, 1988), there are also important processing commonalities. One of the most striking commonalities, identified in both theory and empirical findings, is the overwhelming role of prior knowledge in structuring otherwise "novel" ideas. That is, the conceptual building blocks may vary from science to music to everyday activities, but the tendency to build novel ideas from those domain-specific blocks is the same. In his theories of creative behavior, for instance, Perkins (1981, 1988) argues that creativity primarily consists of harking back to old ideas and reassembling them in some

novel fashion to produce a conceptual combination whose features will resemble the old ideas as well as take on new properties of their own (see, e.g., Gagne & Shoben, 1997; Hampton, 1987, 1997). Ward (1994, 1995) has expressed a similar view in his statement that novel products are rarely entirely novel, insofar as they possess features of the ideas or the products that were used in the creative combination process. Because participants tend to hark back to old ideas, perhaps by applying a systematic search among relevant knowledge structures, their "novel" products tend to be much less original than they otherwise could be (see, e.g., Jansson & Smith, 1991). In a series of previous articles, we have attempted to open a small window on creative cognition by testing participants in a laboratory setting as they perform various generative tasks. We then analyzed their productions for systematic regularities, to assess the way in which prior knowledge influences novel idea generation. Here we extend that line of research to a new domain and a new set of issues. We will briefly describe some of that work, note the relationship to other studies that have been conducted, and describe the psychological questions addressed in the present study.

One task that has been used successfully to study generative cognition is a simple drawing task. The basic paradigm consists of asking participants to draw space creatures to inhabit a distant planet. When participants are given no other information, they tend to incorporate the features of Earth animals into their novel designs. For example, Ward (1994) found that participants' designs uniformly contained the features of bilateral symmetry, sense organs for obtaining information, and appendages for lo-

---

The order of authorship for this article was determined arbitrarily. Appreciation is expressed to Jason L. Hicks for his helpful comments in discussions of this work. We thank Maria Banderas, Martin L. Bink, Rob Brown, and Ted Parsons for their dedicated help in collecting and scoring the data. Correspondence concerning 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).

comoting and manipulating the environment. This tendency to conform to the attributes of Earth animals in designing novel creatures is quite pernicious. Participants who are asked to draw creatures beyond their "wildest imagination" (Ward, 1994, Experiment 3) or to design creatures that are very different from Earth animals (Ward & Sifonis, 1997) nevertheless incorporate these same features into their designs. Clearly, the use of prior knowledge, perhaps as Perkins has expressed it, exerts a strong influence on the production of novel designs.

Using this same drawing task, Marsh, Landau, and Hicks (1996) found that participants do not stray very far from the features of Earth animals. Marsh et al. (1996) coded participants' drawings for every feature that was included in each novel design. We found that, indeed, most creatures locomoted but rarely did so in nonstandard ways by such means as jet propulsion or propeller. Such features could have been easily adopted but were not. We also found that, when participants were shown three examples, all of which contained three common features (e.g., four legs, antennae, and a tail), these three features appeared much more often than they did when participants were not shown any examples at all. Thus, providing examples increased the conformity effect (see, also, Smith, Ward, & Schumacher, 1993, for a similar result).

There are two interesting aspects to this finding that participants' novel creations conform to experimenter-provided examples. First, it provides evidence for the counterintuitive finding that examples do not always improve performance. If divergence and creativity were the goals of a generative task, providing examples decreased performance. Second, because participants were specifically admonished not to copy any of the features used in the examples, the finding suggests that participants have difficulty adhering to such instructions and avoiding the influence of prior knowledge. A similar result was obtained by Smith et al. (1993), who asked participants to diverge as much as possible from the examples. These results are similar to Ward and Sifonis's (1997) finding that participants could not avoid drawing space creatures that resembled Earth animals.

These results, however, bear more than a passing resemblance to our recent work on unconscious plagiarism during idea generation (Marsh, Landau, & Hicks, 1997). In those experiments, we asked participants in small groups to generate ways to improve the university and to generate ways to reduce traffic accidents. When the groups were reassembled 1 week later, we asked each group member individually to generate four new ideas on each topic that had not been given the week before. Thus, participants were specifically admonished to avoid offering an idea that was not novel. Surprisingly, despite that instruction, about 20% of all of the "novel" ideas produced during the second session were plagiarisms from the first session.

Following idea generation during the second session, we administered a source-monitoring test. On this test, participants were presented with all of the old ideas from the first session intermingled with new ideas and were

asked to identify which they had contributed last week, which someone else had offered, and which were new. Plagiarism in this task varied in the low single digits to less than 1%, as compared with the 20% during the generative task. Therefore, source monitoring during generative cognition (idea generation) was much worse than that during nongenerative cognition. The important point to draw from those results is that participants had a lot of information stored in memory that could have been used to comply with the admonition instructions to avoid copying features of the examples (or a complete idea unit in idea generation), but they did not use that information in the process and context of trying to generate a novel product. We have studied and labeled these effects under the umbrella term of cryptomnesia, or unconscious plagiarism (see also Brown & Murphy, 1989; Landau & Marsh, 1997; Marsh & Bower, 1993; Marsh & Landau, 1995).

We are not the only ones to find such regularities in generative cognitive tasks. Rubin, Stoltzfus, and Wall (1991) asked participants to generate novel names for new products (e.g., laundry detergents) and found that participants tended to conform to existing product names. More specifically, participants used the same number of syllables, the same word endings, and the same sort of word stems as those that were quite common for existing exemplars from each of the several categories tested (e.g., analgesics, pastas, etc.). Although Rubin et al. emphasized that their results demonstrated that participants use forms of organization other than semantics (which they do), those results also demonstrate a highly regular cognitive approach to devising a new product name. That approach appears to include the inspection and the use of highly similar and analogous information about the domain of the problem.

The theme running through the studies that we have summarized thus far appears to be that participants will inadvertently use prior knowledge from examples or previously stored concepts, even under specific admonition instructions to avoid doing so. An interesting exception to this general pattern, which partially motivated the present experiments, emerged from a clever set of experiments by Tenpenny, Keriazakos, Lew, and Phelan (1998), who found no evidence for unconscious plagiarism with entirely novel materials. In a typical unconscious plagiarism task, participants are (1) exposed to reasonably familiar information (e.g., known category exemplars) in the context of a generative task, (2) given admonition instructions to avoid using that information, and (3) finally placed once again into the same generative context to determine whether they can avoid using the information they had learned earlier. In contrast, Tenpenny et al. had participants alternate with the experimenter in generating completely unfamiliar nonwords to common categories (e.g., "words" from an extraterrestrial language for various sports, four-legged animals, etc.) and found that participants did not later copy those nonwords that they had been exposed to when asked to generate new nonwords. This result is somewhat surprising, because unconscious plagiarism has been shown to be quite robust when partici-

pants generate real category exemplars (see, e.g., Brown & Halliday, 1991; Brown & Murphy, 1989).

The nonword findings reported by Tenpenny et al. (1998) raise important questions about the influence on creative generation of prior exposure to different types of information and the forms in which that influence is likely to be observed. Do examples only influence generation when they are already familiar (e.g., known category exemplars), thereby allowing an activation of previously stored, coherent representations? Might it be possible to detect a more subtle influence of novel examples by building in coherent structuring properties across multiple examples? Because at least some creativity may occur in largely unexplored domains, the question of whether and how completely novel examples influence subsequent generation is an important one.

One interpretation of the nonword results is that the prior encoding episode may not have fostered any representation (i.e., coherent prior knowledge) that could be used when the generative task was reinstated the second time around. We hypothesized that, if the nonwords possessed common rules that could be learned, participants' novel nonwords might reflect those rules when they were placed in the context of a generative task for which those rules might hold some relevance. That is, participants might not copy whole novel examples, but they might copy newly acquired category properties, even when discouraged from doing so.

Of course, if our hypothesis is incorrect, prior knowledge about completely novel examples would be entirely absent from participants' novel creations. Some support for this alternative prediction was found in a post hoc analysis of accidental regularities contained in the Tenpenny et al. (1998) nonword stimuli. In that analysis, repetitions of orthographic components common to several of the nonwords (e.g., the -on ending) were not used to any greater degree than those orthographic components that were not repeated. Importantly however, Tenpenny et al. did not systematically design materials to contain strong orthographic regularities, and the accidental regularities may not have been prominent enough for an effect to be observed. As is described next, we attempted a similar analysis with stimuli that were designed to contain a highly regular orthographic structure. To the extent that exposure to multiple novel examples establishes a structured representation, whether consciously accessible or completely unreportable, novel creations are expected to reveal the influence of that structure. Thus, our prediction is that participants will be sensitive to the regularities in a set of examples, and, on the basis of our previous work, they should have difficulty in avoiding the use of those features in a generative task, despite strong admonitions to do so.

The general procedure was similar in each of four experiments. In Experiments 1–3, participants learned nonword equivalents paired with six real category exemplars for each of six different categories (e.g., for the category of clothing a *shirt* might be paired with the nonword *beang*). The nonword equivalents were systemati-

cally varied for a given category either to contain specific rules about their orthographic structure or to contain no such structure (a between-subjects manipulation). After learning, participants generated their own nonwords to new exemplars of the same categories under strong admonition instructions to avoid using the features of the nonwords that they had seen for a particular category. In Experiment 1, this generative task occurred after each category was learned (immediate testing), and in Experiment 2, it occurred after all six categories had been learned (delayed testing). In Experiment 3, we tested whether participants could explicitly recall the orthographic regularities that we admonished them not to use (cf. Marsh et al., 1997, on idea generation) and also varied whether the examples were in full view or hidden during generation. In Experiment 4, a different set of categories was used in which the naturally occurring (i.e., real-world) orthographic structure of the exemplars was transferred to different categories and compared with an arbitrary set of features on a set of exemplars.

## EXPERIMENT 1

Experiment 1 was conducted to determine whether participants would use features of nonwords when asked to devise novel nonwords. The participants were taught six nonword exemplars that were paired with real category exemplars (e.g., *dress-ating*). After learning for each category, participants were asked to devise the nonword member of a pair for two new exemplars of the same category (e.g., *shoe-?*). This procedure was repeated five more times with five new categories. The procedure of using nonwords with their own arbitrary regularities along with the admonition instruction allows us to test the generality of our previous findings that participants who perform generative cognitive tasks do not avoid using specific features that they are asked to edit out of their novel creations.

### Method

**Participants.** Ninety-eight University of Georgia undergraduates volunteered to participate in exchange for partial credit toward fulfilling a course research requirement. On the basis of their arrival at the laboratory, participants were randomly assigned to the *consistent*, *inconsistent*, or *control* conditions, as will be described shortly. The number of participants in each group is given in the table of results. The participants were tested in small groups of 3 to 6 participants.

**Materials and Design.** Six categories were selected from the Battig and Montague (1969) norms for use in Experiments 1 and 2. These are listed in Appendix A. For each category, the eight normatively most frequent items were selected for the learning and generation phases of the experiments. Six of the eight exemplars were used in the study portion of this experiment, and the remaining two were used in the generation task in which the participants had to devise their own novel exemplars (the mapping of exemplars to study and test was fixed). These two test items can be obtained from the category norms. Each of the exemplars was paired with a pronounceable nonword, to be shown during the study phase. For the participants assigned to the consistent condition, each nonword exemplar conformed to three orthographic rules relating to the number of syllables, the number of letters, and its ending. As an example,

six common fruit exemplars (*apple, pear, banana, grape, plum, and cherry*) were paired with nonwords that had one syllable, were three letters long, and ended with the letter D. The full set of exemplar-nonword pairs is given in Appendix A, as well as the specific orthographic rules used in constructing the nonwords for each category. The orthographic rules were chosen arbitrarily, but care was taken to ensure that there was no correlation with the English exemplars that comprised the eight normatively most frequent items from the category.

The participants assigned to the inconsistent condition saw the same exemplars and nonwords as participants assigned to the consistent condition. In this condition, however, the nonwords in a given category did not conform to any specific orthographic rules. This manipulation was accomplished, for example, by taking the first nonword from each of the consistent categories and pairing it with each exemplar in the weapons category (e.g., *knife-trenode, gun-ber, bomb-beang*, etc.) and taking the second nonword from each category in the consistent condition and pairing it with each exemplar in the insects category (e.g., *fly-pirlane, wasp-cir, ant-moing*, etc.), and so forth. Thus, each category in the inconsistent condition had one nonword from each of the six sets of rules that were used in the consistent condition. In this fashion, the six exemplars for a given category in the inconsistent condition conformed to all 13 orthographic manipulations (i.e., one, two, or three syllables, three, five, seven, or eight letters, and endings of E, R, NG, Y, T, and D). Our intuition is that there is no obvious orthographic rule in the inconsistent stimuli, but we address this point later in the Results and Discussion section.

**Procedure.** Prior to learning, the participants heard the following cover story, which the experimenter read aloud.

Our laboratory is developing a computer program that generates novel words to replace common, everyday words. So, if we type the word "school" into our computer program, our computer takes that word and generates a new word, for example, "kiflinchet." Although we are quite pleased with what the computer is generating, we would like to add a human element to our program that is currently missing. Therefore, we want to see what sorts of new words you will generate when asked to come up with brand new words. For each of six categories, we will show you six examples of what our computer has generated and then ask you to generate two more.

For each of the categories, the six exemplar-nonword pairs were presented one every 5 sec on an overhead projector. As each pair was presented, the experimenter pronounced it aloud. After all six pairs had been presented for a given category, the participants were shown two real exemplars (e.g., *tank* and *spear* for weapons) and asked to generate a novel word for each. Special booklets had been prepared that had the two new exemplars and a blank space for providing the novel words that the participants were to devise. Before offering their new words (for each of the six categories), the participants were admonished not to copy any aspect of the examples, and they were also reminded not to use words from a foreign language. These instructions were as follows:

Below the category heading there are two category members followed by a blank space. Your task is to create a brand new word for each member. Keep in mind, however, that you should NOT copy or use any aspects of the examples that we showed you. Also please avoid using words from a foreign language when creating your new words.

One minute was given for this generative activity (which was ample time for everyone to complete the task). Following generation, the next category was presented in a manner identical to that just described. The order of the categories was completely random for each group of participants tested. The procedure was, therefore, identical in the consistent and inconsistent conditions, and the only thing that differed was the exemplars that comprised the nonwords in each category. The participants in the control condition had the cover story read to them (with a modified final sentence), just as in the other two conditions, but they did not see any exemplar-nonword

pairs. The same booklet as that used in the inconsistent and consistent conditions was used to collect control participants' novel words.

**Results and Discussion**

Unless otherwise noted, the probability of a Type I error in the statistical analyses does not exceed 5% in this experiment and the ones that follow. To minimize the probability of an error in scoring, two raters unaware of the aims of the experiments scored the data. Their results were almost entirely consistent with one another, and in the very few cases where their scoring diverged, the discrepancy was resolved in conference (legibility of handwriting was the issue in these very few cases).

**Category-specific measure.** In a first analysis of the data, the exemplars were scored for whether they contained any of the rules used to construct the nonwords shown in the consistent condition for that particular category. For example, the participants in all three conditions had their novel weapons scored for the proportion of exemplars that were two syllables long, were seven letters in length, or ended with an E. Likewise, all conditions had their novel insects scored for the proportion that were one syllable, were three letters in length, or ended with an R (see Appendix A). We call this a *category-specific* measure of conformity to the examples. The average proportion of the participants' novel exemplars conforming to the rules of the consistent condition is given in the upper half of Table 1. In that table, the data are provided for each orthographic rule, and the average of the three rules is given as the overall conformity score reported in the final column of Table 1.

The overall conformity score showed that the participants who were provided consistent orthographic rules conformed to those rules to a greater degree than did the participants who either learned a set of inconsistent rules or learned nothing at all prior to generating (i.e., the control group). That overall result was statistically significant [ $F(2,95) = 23.8, MS_e = 0.02$ ]. The equivalent performance of the control and inconsistent groups (and greater conformity in the consistent group) was preserved over each of the three individuals rules as well [smallest  $F(2,95) = 11.2, MS_e = 0.03$ ]. Post hoc analyses (by Scheffé tests) confirmed the obvious pattern that conformity was greater in the consistent group than in either of the other two groups (minimum difference needed was .09). This pattern of results suggests that participants who

**Table 1**  
Proportion of Novel Words Conforming to the Orthographic Rules of the Nonwords Shown in Experiment 1

Condition	N	Number of Letters	Number of Syllables	Word Ending	Overall Conformity
Category-Specific Measure					
Consistent	30	.33	.53	.26	.37
Inconsistent	34	.12	.34	.08	.18
Control	34	.10	.30	.08	.16
Experiment-Wide Measure					
Consistent		.59	.91	.52	.67
Inconsistent		.50	.94	.47	.64
Control		.49	.94	.35	.59

are shown a set of consistent exemplars learn and retain the structure of those exemplars. When asked to create novel exemplars for items from the same category, this prior knowledge appears to be used as a basis for generating something new relating to that category, even though we admonished them against using the features of the examples.

**Experiment-wide measure.** The fact that the inconsistent group appeared to conform to an equally small degree as the participants shown nothing at all does not necessarily imply that they learned nothing from the examples they were exposed to. Rather, the absence of an observable impact of the examples may be the result of the stringent, category-specific scoring based on the rules shown to the consistent group. The inconsistent group was exposed to the same rules experiment wide as the consistent group, but specific rules were not linked uniquely to specific categories. Therefore, in another scoring of the data for each of the three groups, participants' novel creations were scored as to whether the number of letters, number of syllables, and endings matched any of the 13 rules used across all six categories (these rules are listed in the Method section of this experiment). We call this an *experiment-wide* measure of conformity to the examples, and this scoring is provided in the lower half of Table 1. As can be seen immediately, almost every participant at every opportunity exclusively generated a one-, two-, or three-syllable nonword, and this was true of all three conditions [ $F(2,95) < 1$ ,  $MS_e = 0.01$ ,  $p > .10$ ]. There was, however, a difference among the three conditions in the overall conformity score [ $F(2,95) = 3.4$ ,  $MS_e = 0.01$ ]. The consistent and the inconsistent groups had greater conformity than the control group, and even the smaller of the two differences, that between the inconsistent and control groups, was significant [ $t(95) = 2.33$ ,  $SE = 0.03$ ]. The equivalent performance of the consistent and inconsistent groups was driven by conformity to the word endings. Post hoc tests showed that the inconsistent group (and their difference from the control group) appears to be driven by higher conformity to word endings than in the control group [ $t(95) = 2.53$ ,  $SE = 0.02$ ]. The same pattern, however, was not observed with the number of letters of newly devised nonwords.

The results from this experiment suggest that participants who are engaged in a generative task find it difficult to avoid using information that was recently experienced. The same was true with the example space creatures provided in the Marsh et al. (1996) and Smith et al. (1993) studies, as well as with the ideas in the Marsh et al. (1997) study. In addition, Ward's (1994) results suggest the same phenomenon of not being able to avoid the use of long-standing knowledge about Earth animals. The equivalent experiment-wise conformity of the two groups shown examples also suggests that participants in the inconsistent condition may use word endings they had learned from prior categories. That is, by an experiment-wide measure, conformity to word endings was similar in both the consistent and the inconsistent conditions. Word endings may be particularly salient, either in the seman-

tic meaning they impart in English or because certain categories have similar endings in the English language and participants are sensitive to that fact (Rubin et al., 1991). Converging evidence for the second of these two possibilities was found by Marsh et al. (1996, Experiment 2). In that study, participants conformed more to examples whose features tended to co-occur in the real world than to examples whose features did not. That evidence is also consistent with Murphy and Allopenna's (1994) finding that participants have a more difficult time learning about categories whose structure is inconsistent with their real-world knowledge, as compared with cases in which that information is consistent.

## EXPERIMENT 2

The goal of Experiment 2 was to explore the very subtle finding that even the inconsistent group retained information about the orthographic rules by an experiment-wide measure, despite not having learned them consistently (i.e., coherently grouped together by category). Our intuition is that some aspects of learning the rules could be equivalent in the consistent and inconsistent groups by the time the last category has been presented. That is, by the time all the categories have been presented, the participants in the consistent and inconsistent groups have been exposed equally to the full set of rules that govern the novel exemplars used in the experiment, although only the former group has been exposed to the category-specific rules. Therefore, they might be expected to have equivalent performance on the basis of the experiment-wide measure. To determine whether this was true, in Experiment 2, both the consistent and the inconsistent groups generated their novel items after all exemplar-nonword pairs had been shown for each of the six categories. How the groups might perform relative to one another on the category-specific measure after exposure to all the categories is less clear.

On the one hand, the participants in the consistent group may be able to remember aspects of the exemplar-nonword pairs that they saw earlier for a given category and subsequently use that information on a delayed test, despite admonitions to avoid doing so. If this occurred, their conformity to the specific rules would still be greater than that exhibited by the inconsistent group. On the other hand, the consistent group may perform very much like the inconsistent group on a delayed test because the knowledge acquired about each category becomes fairly undifferentiated after seeing six categories of nonwords. In this case, learning in the consistent group would resemble the inconsistently learned set of orthographic rules. This alternative prediction is supported by the very large literature on the rapid loss of surface form (e.g., Brewer & Nakamura, 1984; Wallace & Rubin, 1991).

Under either prediction, it would be informative to know whether participants had explicit recollection for the nonword items. Marsh et al. (1997) found very strong evidence that participants possessed a great deal of recollective information that could have been used to avoid

copying the aspects of the encoding experience but was not used during the generative task. In order to assess recollection, a cued recall test was administered after the generation phase in this experiment.

### Method

**Participants.** Ninety-two University of Georgia undergraduates from the same pool as that in Experiment 1 volunteered in exchange for partial course research credit. The participants, none of whom had served in Experiment 1, were randomly assigned to the same three encoding manipulations as those used in Experiment 1. Group sizes are displayed in the table of results.

**Materials and Procedure.** The materials were identical to those in Experiment 1 (see Appendix A). Three modifications were made to the procedure. First, rather than having novel category exemplars generated immediately after exposure to the exemplar–nonword pairs, all six of the categories were presented first, and then the participants generated their novel nonwords. Therefore, all testing occurred after a delay. Second, instead of generating two new words, as in Experiment 1, participants generated 4 new items per category. The 9th and 10th items from the Battig and Montague (1969) norms were used for these 2 additional items. This change was made to increase the sensitivity of the experiment to any differences that were present after the delay of learning all six categories. Third, following the generation task, a cued recall test was administered to the participants in the consistent and inconsistent conditions to determine whether they could remember any of the experimenter-provided exemplars (i.e., word–nonword pairs). On this test, the six categories were presented on paper along with the six English words. Beside each exemplar was a blank space for the participants to write in their recall of the nonword studied during the encoding phase. We required that the participants spend a minimum of 5 min on the cued recall task.

### Results and Discussion

The critical results are set forth in Table 2. The major difference between this experiment and Experiment 1 was the delay between learning and testing for the consistent and inconsistent groups. As is evident in Table 2, the greatest difference from Experiment 1 for the category-specific measure was the large decrease in conformity evidenced in the consistent learning group (i.e., about one third less conformity). In terms of the overall conformity effect, however, there were significant differences among the conditions [ $F(2,89) = 4.9, MS_e = 0.01$ ]. Post hoc tests revealed that only the 4% difference between the consistent and the control groups was significant on this overall measure [ $t(89) = 2.98, SE = 0.01$ ]. This overall conformity effect was driven by the individual orthographic

rule of the number of letters, for which the consistent and control group differ as well [ $t(89) = 3.2, SE = 0.02$ ]. Therefore, these results suggest, once again, that prior learning affected later generative activities, although, comparing across experiments, the expression of that knowledge is larger when testing occurs immediately rather than after a delay. Clearly, the overall effect size is rather small. It is impressive, however, that the conformity occurred at all in the consistent condition with this category-specific measure, because the nonword stimuli are virtually devoid of meaning and memory for orthography is far worse than that for semantics (see Alba & Hasher, 1983, or Gernsbacher, 1985). Even after exposure to 36 novel nonwords across six categories, the participants maintained at least some information about the appropriate category assignments of the specific orthographic rules.

In terms of the experiment-wide measure of conformity, the results replicated those for Experiment 1 (see the lower half of Table 2). Overall conformity was reliably different among the three groups [ $F(2,89) = 3.62, MS_e = 0.01$ ]. As in Experiment 1, the level of conformity in the inconsistent group fell in between those of the consistent and control groups, but only those latter two groups differed statistically [ $t(89) = 2.69, SE = 0.02$ ]. The analysis of the word endings wholly replicated Experiment 1 for this experiment-wide measure. The consistent and the inconsistent group conformed much more to the endings of the nonword examples than did the participants in the control condition, who saw no examples [ $F(2,89) = 3.84, MS_e = 0.03$ ]. None of the groups differed on either of the remaining two rules [largest  $F(2,89) = 1.8, MS_e = 0.01, p > .15$ ]. In this experiment and Experiment 1, there is a practical limit on the number of syllables and the number of letters participants will use. For example, participants may never generate a novel nonword with more than two or three syllables. Thus, our examples may have exhausted the realistic possible variations, and this may be why we failed to find any differences in these two attributes.

Performance on the cued recall test was uniformly poor. In fact, virtually everyone reported that it was impossible to recall correctly more than about one of the nonwords that were learned earlier. In the consistent condition, mean recall of the six items in a category was 2%, and in the inconsistent condition it was 2.5%. These exceptionally poor overall levels of recall represent floor effects. In retrospect, our choice of measurement of conscious recollection might have been inferior to simply asking participants to write down the number of syllables, common endings, and the number of letters. That is, participants may well have information available about the rules that govern the domain without being able to recall any exact instance from that category. If participants have explicit knowledge of the rules and use them anyway, despite our admonitions, this would be strong evidence that participants performing generative cognitive tasks do not utilize all of the information they have available to comply with the task demands. In the next experiment, we assess conscious recollection of the rules directly.

**Table 2**  
Proportion of Novel Words Conforming to the Orthographic Rules of the Nonwords Shown in Experiment 2

Condition	<i>N</i>	Number of Letters	Number of Syllables	Word Ending	Overall Conformity
Category-Specific Measure					
Consistent	30	.17	.37	.09	.21
Inconsistent	31	.12	.33	.08	.18
Control	31	.10	.33	.07	.17
Experiment-Wide Measure					
Consistent		.55	.99	.43	.66
Inconsistent		.52	.94	.43	.63
Control		.51	.97	.33	.60

### EXPERIMENT 3

The purpose of this next experiment was to assess conscious recollection by asking participants to provide the orthographic rules after generating all of their exemplars to all six categories.<sup>1</sup> We used only the consistent condition from Experiment 1, in order to test a strong hypothesis from the Marsh et al. (1997) study. That hypothesis is that participants will fail to use relevant information that they have stored to avoid conformity (i.e., unconscious plagiarism). A strong test of this hypothesis would be to leave the word–nonword pairs in full view when the participants were generating their novel nonwords. If there was any ambiguity about whether a feature should be avoided, all the participants would have to do is glance up at the screen in order to ascertain whether it was a component in the examples. In Experiments 1 and 2, the participants saw each pair for 5 sec as the experimenter read them aloud. In this experiment, we compared such a hidden condition with an in-view condition in which the six nonwords for each category were available for consultation during the generation of novel nonwords.

In addition to providing the strongest test of whether participants use information that they have stored, the in-view versus hidden manipulation acts as an encoding manipulation in terms of exposure as well. In the in-view condition, the participants will have all of the word–nonword pairs presented for a total of 1 min 30 sec (original encoding plus the 1-min generation period). Therefore, when we assess recall of the orthographic rules, explicit memory should be better in the in-view condition than in the hidden condition. In contrast to this explicit measure of performance, conformity to the rules could be either higher in the hidden condition or equivalent. If conformity is higher in the hidden condition, that would suggest that the participants in the in-view condition either checked the exemplars or consulted their generally better memories, to be sure they were avoiding features. Although we could not disambiguate by which means they were avoiding unconscious plagiarism, the result would be the same: less conformity to the experimenter-provided examples. In contrast, if participants do not use information that they have available, to avoid unconscious plagiarism, as Marsh et al. (1997) have strongly argued, recall of the rules might be different, but conformity would be equivalent. If this latter prediction came to pass, it would demonstrate one important feature of generative cognition—namely, that participants engrossed in generative cognitive tasks do not consider the origin of the components of their solutions to a sufficient degree and, therefore, do not edit out components that should be avoided.

#### Method

**Participants.** Forty-four volunteers participated in exchange for partial credit toward meeting a research requirement. None had participated previously. Approximately half were assigned to a *hidden* condition, and the remainder were assigned to an *in-view* condition.

**Procedure.** The procedure generally followed that of the consistent condition in Experiment 1. The participants were shown six word–nonword pairs for each of six categories. These were the consistent stimuli used in Experiments 1 and 2. For the hidden condition, each pair was presented for 5 sec, as was done in the previous experiments. For the in-view condition, the cardboard mask that was used to reveal the pairs of stimuli was replaced with a solid mask that was simply drawn down the transparency to reveal each successive pair. Thus, the pairs shown first remained on the overhead. These were also left fully in view when the participants generated their novel nonwords for each of two new category exemplars. After the 1-min generation period, the next category was shown in an identical fashion, depending on the condition being tested. After this study–test sequence on all six categories, we tested the participants' recollection of the rules by giving them a response sheet that had all the English equivalents for each of the categories and places to write in the orthographic rules. The instructions, printed on the page and read by the experimenter, were as follows:

On this page we have listed the six categories and their corresponding English exemplars that you saw earlier. Below each category three questions appear. Your task is to think back to the nonwords that were created by the computer and to fill in the following information for each category that describes the nonwords: the number of syllables, the number of letters, and the common ending (last letter or letters). If you are unsure about some of the answers, then please do your best to fill in every space, even if you have to guess.

All other study and admonition instructions were the same as those reported in Experiment 1.

#### Results and Discussion

The results are set forth in Table 3. As can be seen in that table for the category-specific measure, overall conformity to the orthographic rules did not differ for the hidden and the in-view conditions [ $t(42) < 1.0, p > .50$ ], nor did it differ for any of the rules individually [largest  $t(42) = 1.27, p > .20$ ]. The same generally small sampling differences were observed with the experiment-wide measure. Overall, conformity did not differ between the hidden and the in-view conditions [ $t(42) < 1.0, p > .50$ ], nor did conformity to any of the individual rules differ in the conditions [largest  $t(42) = 1.03, p > .30$ ]. These results strongly suggest that even when participants have the word–nonword exemplars in full view, they do not efficiently use them, in order to avoid plagiarizing aspects of the exemplars. These results are identical to Marsh et al.'s (1997) finding that participants inadvertently plagiarized ideas during a generative task, despite later being able to identify that they heard the ideas before and to specify correctly the source of these old ideas. Together, there seems to be convergent evidence that, while engaged in generative tasks, participants fail to consider the source(s) of the components of their novel production.

Performance on the explicit measure of recollection taken at the end of the experimental sequence is shown in the last two rows of Table 3. As can be seen there, there was an overall significant difference in recollection. The participants tested in the in-view condition could remember more of the rules than could those tested in the hidden condition [ $t(42) = 2.39$ ]. Although this effect was observed with each of the rules individually, as in the previous experiments, only a statistical difference was observed in the

**Table 3**  
**Proportion of Novel Words Conforming to the**  
**Orthographic Rules of the Nonwords Shown in Experiment 3**

Condition	<i>N</i>	Number of Letters	Number of Syllables	Word Ending	Overall Conformity
Category-Specific Measure					
Hidden	23	.25	.46	.12	.28
In-view	21	.24	.46	.17	.29
Experiment-Wide Measure					
Hidden		.56	.99	.47	.68
In-view		.60	.97	.50	.69
Recollection of Rules					
Hidden		.36	.59	.32	.43
In-view		.45	.68	.46	.53

number of word endings [ $t(42) = 2.10$ ]. There were, however, differences in the marginal range for number of letters [ $t(42) = 1.56, p = .12$ ] and number of syllables [ $t(42) = 1.50, p = .14$ ]. Although the overall differences between the category-specific measure and the recall measure might represent a measure of the degree to which participants could avoid plagiarism, the important point is that recollection was better for the in-view group but their use of the rules was equivalent to the hidden group. Thus, despite the fact that the in-view participants had more information that they could use to avoid copying the features of the experimenter-provided examples, they did not use it. Although there are probably generative tasks and conditions that enable participants to utilize better the information that they have available, they do not seem able (or willing) to do so in the several tasks that we have tested.

#### EXPERIMENT 4

There were several motivations for conducting this last experiment on generating novel nonwords. First, we wanted to compare participants' conformity to an arbitrary orthographic rule with their conformity to a "rule" that occurred naturally in the English language. In effect, we wanted to examine the influence of a newly established prior knowledge, relative to familiar, well-established prior knowledge. Previous research using creative generation paradigms has revealed effects of recently presented materials (e.g., Marsh et al., 1996; Smith et al., 1993), as well as those of long-standing knowledge (Ward, 1994). Perhaps participants are better able to avoid long-standing knowledge, or perhaps they are less able to do so. In this next experiment, we examined the influence of both types of information within the same experiment. To accomplish this, we placed arbitrary but consistent word endings (which was arguably the strongest orthographic rule learned in Experiments 1–3) on nonwords in categories that naturally have other uniform endings (e.g., an arbitrary -r ending on *chir* for a category such as pastas, whose exemplars normally end in -i and -a; see Rubin et al., 1991). Thus, the participants were not learning

exemplar–nonword pairs; they merely studied a set of six novel nonword exemplars for each category.

Second, for generality, we wanted to ascertain whether participants' sensitivity would be different to arbitrary word endings that we made up, as compared with endings that do naturally occur in English but that occur for other categories than the ones of interest. To accomplish this, naturally occurring endings for a set of categories were mismatched to form new nonwords (i.e., what were natural endings for Category A were placed on the word stems for Category B, etc.). Third, we wanted to examine the differences between immediate and delayed testing in a single experiment. In doing so, however, we also wanted to reduce the amount of information that participants learned between immediate and delayed testing and, as a consequence, reduced the number of categories to three. The rationale was that, with less information being learned, perhaps a stronger effect could be seen after a delay (cf. Experiment 2). Accomplishing these goals required four different encoding conditions, as is described next.

#### Method

**Participants.** Ninety-eight University of Georgia undergraduates volunteered in exchange for partial course research credit. The participants, none of whom had served in the previous experiments, were randomly assigned to the *arbitrary* ( $n = 24$ ), *mismatched* ( $n = 19$ ), *natural* ( $n = 33$ ), or *control* condition ( $n = 22$ ), as will be described shortly. The participants were tested in small groups of from 3 to 6 participants.

**Materials.** Three categories (analgesics, elements, and pastas) were chosen from the set that was originally used by Rubin et al. (1991). In the natural condition, six naturally occurring exemplars for each of the three categories were taken from that earlier study (e.g., *spaghetti* for pastas, etc.). The stimuli are given in Appendix B. In the mismatched condition, the typical endings for a category were replaced with the typical endings for another category. Thus, as can be seen in Appendix B, pastas ended in -on or -ium (from elements), the elements ended with -ol or -in (from analgesics), and the analgesics ended with -i or -a (from pastas). These substitutions themselves were determined with no a priori reasoning and did not change throughout the experiment (note that there are only two ways to mismatch three category endings). In the arbitrary condition, both the stem and the endings were arbitrary, but the endings were consistent, as in the natural and mismatched conditions of this experiment and the consistent condition of Experiments 1 and 2. The control group did not see any examples.

**Procedure.** The procedure for this experiment was virtually identical to that used in Experiment 1. The basic cover story used in Experiment 1 (modified slightly to conform to the stimuli and procedures used in this experiment) was used again. The participants studied six category "exemplars" on an overhead (which the experimenter read aloud) for 5 sec each (cf. the hidden condition of Experiment 3). In the arbitrary and mismatched conditions, these appeared to the participants as nonwords similar to those they were told they would have to generate. In the natural condition, they were read for what they were: real exemplars of the category from which participants would have to generate new nonwords. The only difference from Experiment 1 in the encoding experience, besides the stimulus conditions just described, was the fact that participants did not study exemplar–nonword pairs; rather, they studied single exemplars. Immediately after seeing the six exemplars, the partici-



pants were asked to devise three brand new nonwords under the admonition not to copy any aspect of the examples. Two minutes were given to create the three new exemplars. The remaining two categories were tested in an identical fashion. The order of the categories was completely random for each group of participants. After this immediate generation of three exemplars per category, the participants were given 6 min at the end of the experiment to create nine additional exemplars (three per category). In both the immediate and the delayed tests, the participants wrote their new nonwords in a booklet that had three blank spaces below a category label. On both tests, the categories were given in the same sequential order as that in which they were studied. Because of this sequential nature of the testing, on average approximately 4 min elapsed between the immediate and the delayed tests. The participants were, as in Experiments 1 and 2, required to provide a response in each blank.

### Results and Discussion

For simplicity of exposition, the data were scored and are reported for each of the four encoding groups according to the category-specific metric used for each of the three types of endings. Therefore, the columns of Table 4 correspond to encoding groups, and the rows correspond to the various types of orthographic rules tested in the three groups of participants who saw exemplars. The data are unambiguous and are considered in turn for the immediate and delayed testing. When scored for the arbitrary endings (i.e., first row of Table 4), the greatest conformity was observed in the participants who were exposed to those arbitrary endings [ $F(3,94) = 12.4$ ,  $MS_e = 0.01$ ]. Likewise, when scored for the mismatched endings (e.g., the proportion of new pastas with -on and -ium endings), only the participants in the mismatched group tended to conform to the new endings [ $F(3,94) = 11.5$ ,  $MS_e = 0.01$ ]. Interestingly, the 8.19% conformity was similar in magnitude to the conformity (8.33%) in the arbitrary condition's scoring for the arbitrary features. Evidently, these naturally occurring endings that occurred in the contexts of different categories appeared to be just as arbitrary as the ones chosen to be arbitrary.

The final comparison is for the scoring of the novel creations to their natural endings. As is obvious from that row of Table 4, the largest conformity effects in this experiment were observed under that scoring, and the

differences among the four encoding conditions were significant [ $F(3,94) = 35.5$ ,  $MS_e = 0.01$ ]. Reminding participants of existing category members increased their propensity to use the natural word endings from that category, as compared with participants in the control group, who received no such reminding [ $t(94) = 5.8$ ,  $SE = 0.03$ ]. That effect is quite striking, despite being thrice admonished to avoid copying any aspects of the examples shown to them. Performance in the control group, who saw no examples (.17), is entirely consistent with Rubin et al.'s (1991) finding that participants are sensitive to and use a naturally occurring orthographic structure when one is available. Equally as important, showing participants arbitrary or mismatched endings decreased their propensity to include the natural endings (11.2%), as compared with the control (17.4%) condition [contrast,  $t(94) = 2.50$ ,  $SE = 0.03$ ]. On the whole, we believe that these data demonstrate that participants use prior knowledge, either recently experienced or long-standing, acquired during learning in order to devise novel category members, despite strong admonitions to avoid using such knowledge.

Interpreting the data from the delayed generation portion of this experiment is slightly complicated by the fact that the participants had generated three exemplars earlier to each category during immediate generation. Nevertheless, the pattern of effects neatly replicated the data from immediate generation. As can be seen by comparing the upper and lower halves of Table 4, the only significant difference in the 12 proportions comprising each half of the table was isolated to scoring the arbitrary condition for the arbitrary endings (before Bonferroni correction for multiple comparisons). Specifically, the only change from immediate to delayed testing was a decline in specific conformity to arbitrary endings by the participants exposed to those endings. Put slightly differently, even after a delay, the participants exposed to mismatched words endings still used that mismatched information directly in their novel creations, and participants exposed to either arbitrary or mismatched endings showed an indirect influence of that information, as is evidenced by a decrease in their propensity to include natural endings when compared with participants in the control group, who saw no examples. We now turn to a discussion of what these results, and those from Experiments 1–3, have to say about participants' performance on a generative cognitive task.

### GENERAL DISCUSSION

The four experiments conducted in this study were undertaken to further elucidate the process of how participants might go about solving generative cognitive tasks. Our hypothesis was that participants who were exposed to information that is irrelevant to solving an open-ended task would use that information, despite admonitions to avoid doing so. The empirical evidence from Experiments 1–4 supports that claim. In each experiment, participants

**Table 4**  
**Proportion of Novel Words Scored for (Rows of the Table)**  
**Conformity to the Word Endings Used on the Arbitrary,**  
**Mismatched, and Natural Stimuli by Experimental Condition**  
**(Columns of the Table) in Experiment 4**

Ending Scored For	Experimental Condition Tested			
	Arbitrary	Mismatched	Natural	Control
Immediate Testing				
Arbitrary	.0833	.0205	.0067	.0076
Mismatched	.0069	.0819	.0084	.0278
Natural	.0741	.1491	.3266	.1742
Delayed Testing				
Arbitrary	.0347	.0088	.0101	.0126
Mismatched	.0162	.0702	.0219	.0126
Natural	.0764	.1404	.3131	.1818

were exposed to example nonwords that shared certain attributes. When asked to generate new nonwords, regularities that were inherent to the examples were found in the participants' novel productions. Those regularities were strongest when the participants were asked to generate immediately after exposure (Experiment 1), but were attenuated when generation was delayed (Experiment 2). These regularities appear to be a rather general phenomenon in generative tasks, because participants used both newly acquired information and long-standing knowledge about the orthographic structure of existing categories (Experiment 4; cf. Rubin et al., 1991). They did so, however, despite having some access to conscious knowledge of the rules (Experiment 3).

The patterns of results found in this study and others from our laboratories (e.g., Marsh et al., 1996; Smith et al., 1993; Ward, 1994; Ward & Sifonis, 1997) appear to converge on an important point about generative cognition. That point concerns how participants approach the task of generating something novel. Participants appear to take as their starting point relevant prior knowledge, which Perkins (1981, 1988) called a harking back to old ideas. As participants consider ways to assemble that prior knowledge into novel forms, they do not consider the task constraints to avoid the use of some of that knowledge. It is not the case that participants do not have access to the relevant knowledge that is to be avoided. Experiment 3 demonstrated that the participants in the in-view condition had more knowledge of the rules that were to be avoided but conformed to the experimenter-provided examples to the same extent as the participants in the hidden condition, who had less knowledge of the rules. That result suggests that the task demands of generating novel material may cause participants to inadvertently include information in their novel products, because, in other circumstances (e.g., a recall test or source-monitoring test), these same participants can demonstrate that they have knowledge of what is to be avoided.

One interpretation of these results is that participants have misunderstood the task instructions or have simply ignored the admonition instructions altogether. Either explanation would have produced the pattern of results that we have found. Direct manipulations of instructions with similar generative tasks demonstrate that performance does change with instructions that would seem to alleviate such concerns, at least to some degree. For example, Marsh et al. (1997, Experiment 4) manipulated instructions to participants, to avoid copying the prior knowledge acquired earlier, with both strict and lenient instructions. The result was a significant reduction in conformity to that knowledge in the stringent condition, but there was still a substantial conformity effect. Likewise, Smith et al. (1993) manipulated instructions as well and found that participants were sensitive to admonition versus conform instructions. These data argue that participants attempt to comply with task demands but nevertheless fail to some degree to edit out of their novel productions prior knowledge that should not be expressed. Marsh et al. (1997) have claimed that the error

represents a source-monitoring error. They argued that participants fail to consider the source of the components of their novel products, as is required by the task. The failure appears to emanate from the fact that participants are cognitively engrossed in the task at hand—namely, generating a novel entity. Just as source-monitoring performance depends on the context in which it occurs (see Johnson, Hashtroudi, & Lindsay, 1993), output editing is not as efficient as it could be in generative tasks, because participants are cognitively engaged in creating something novel.

The conformity and inadvertent plagiarism effects that we have found occur because participants working on creative tasks fail to engage in the systematic decision processes specified by source-monitoring framework (cf. Jacoby, Kelley, Brown, & Jasechko, 1989). There is, however, the question of why the knowledge comes to mind in the first place. We have implicitly argued that it is consciously retrieved in the search for old components to recombine in novel ways. Although this might be true, it could also come to mind inadvertently as an involuntary retrieval of information. In a series of seven experiments, Smith and Tindell (1997) exposed participants to words (e.g., *allergy*). Later, they asked participants to solve orthographically similar word fragments (e.g., *analogy*). Compared with participants who were not primed, participants exposed to the orthographically similar words failed to solve the fragments to a much greater degree. Smith and Tindell demonstrated that the involuntary retrieval of the material blocked the successful search for an answer. Thus, some of the orthographic regularities that participants used might have come to mind involuntarily, and, without applying systematic decision criteria, participants are likely to commit a source-monitoring error during generative cognition.

Appealing to the source-monitoring framework has the additional appeal that it can shed additional light on Ward's (1994) original results. In that report, participants tended to draw space creatures that consisted of features from Earth animals. We have already noted that features such as propellers, wheels, string, and others from other domains would have been just as valid, and perhaps more creative, if they were included on alien beings. They were not. Generative cognition appears to begin with participants deciding what the most relevant source of information is that pertains to the generative task at hand. That source, or domain, then appears to be searched for elements that might be relevant to devising something novel. Searching within a domain constrains the search but also has the ill-fated consequence that participants are not as creative as they could be. For example, most novel space creatures are bilaterally symmetric and contain appendages and sense organs. Moreover, searching for elements within a domain, combined with a failure to output monitor the features (i.e., monitor source), increases the probability that the novel product contains elements that should not be included. This line of argumentation, which most assuredly awaits further experimentation, suggests that when artisans or scientists devise novel en-

tities, those entities contain plagiarized aspects of which the creator might be entirely unaware.

The results from Experiments 1–4 may seem to be at odds with Tenpenny et al.'s (1998) finding that participants did not plagiarize or otherwise conform to nonwords. In actual fact, the two sets of results complement each other rather nicely. Tenpenny et al.'s materials did not systematically include specific rules that could be learned, similar to the inconsistent groups in Experiments 1 and 2. Without a consistent and strong rule, participants did not express that earlier learning in their generative products. Importantly, the Tenpenny et al. results demonstrate that entire nonwords do not get used in later generative tasks. That finding supports our intuition that participants cannot necessarily recollect the examples in their entirety (Experiment 2) but that consistent aspects of examples (if they are there to be learned) will be elicited in novel products that participants generate.

Of course, it is true that the effect sizes demonstrating this influence on generative activities were not numerically large in any of the experiments, and especially so in Experiment 2 after a delay. Nevertheless, those same effects were remarkably consistent across the four experiments with different materials and slightly different procedures. In assessing the size of these effects, we note that we purposely chose materials that were inherently devoid of semantic meaning. Thus, a priori we expected somewhat small conformity to the nonword exemplars to which participants were exposed. After all, surface structure is lost far more rapidly than is semantic meaning (see, e.g., Gernsbacher, 1985). The decision to use nonwords was motivated in part by our thesis that participants are quite sensitive to prior knowledge and approach generative tasks by searching memory for the most relevant representations and drawing upon that information to solve the task at hand. In previous studies using Ward's (1994) drawing task, those representations appear to be Earth animals or, when participants are shown examples, the examples themselves (Marsh et al., 1996).

We note, in closing, that the potential dependence on prior knowledge to arrive at solutions to open-ended generative problems has important theoretical implications for creativity in general. Participants who possess small and constrained amounts of prior knowledge for particular topics may be less likely to generate truly novel solutions for those topics because their search is more constrained for old elements to recombine. The converse would also seem to be true. Participants who possess large and unconstrained amounts of prior knowledge, or whose boundaries for topics overlap widely, would seem to be able to generate both more and better quality solutions. Likewise, the same might be true of topics that are small and constrained across all participants, as opposed to those that are less constrained. Although generative problems could be solved in a variety of ways, there is mounting evidence (Marsh et al., 1996, 1997; Ward, 1994; Ward & Sifonis, 1997) that one cognitive universal may be the bringing of relevant prior knowledge, both overtly and

more inadvertently, to bear as a first step in the process of arriving at a novel solution.

## REFERENCES

- ALBA, J. W., & HASHER, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*, 203-231.
- BATTIG, W. F., & MONTAGUE, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs*, *80* (3, Pt. 2).
- BREWER, W. F., & NAKAMURA, G. V. (1984). The nature and function of schemas. In S. R. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 119-160). Hillsdale, NJ: Erlbaum.
- 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.
- GAGNE, C., & SHOEN, E. J. (1997). The influence of thematic relations on the comprehension of nonpredicating combinations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *23*, 71-87.
- GERNSBACHER, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology*, *17*, 324-363.
- GREENO, J. G., & SIMON, H. A. (1988). Problem solving and reasoning. In R. C. Atkinson, H. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens's Handbook of experimental psychology* (Vol. 2, pp. 589-674). New York: Wiley.
- HAMPTON, J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition*, *15*, 55-71.
- HAMPTON, J. A. (1997). Conceptual combination: Conjunction and negation of natural concepts. *Memory & Cognition*, *25*, 888-909.
- 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.
- JOHNSON, M. K., HASHTROUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3-28.
- LANDAU, J. D., & MARSH, R. L. (1997). Monitoring source in an unconscious plagiarism paradigm. *Psychonomic Bulletin & Review*, *4*, 265-270.
- 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.
- MARSH, R. L., LANDAU, J. D., & HICKS, J. L. (1996). How examples may (and may not) constrain creativity. *Memory & Cognition*, *24*, 669-680.
- MARSH, R. L., LANDAU, J. D., & HICKS, J. L. (1997). The contribution of inadequate source monitoring to unconscious plagiarism during idea generation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *23*, 886-897.
- 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.
- PERKINS, D. N. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.
- PERKINS, D. N. (1988). Creativity and the quest for a mechanism. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (309-336). New York: Cambridge University Press.
- RUBIN, D. C., STOLTZEUS, E. R., & WALL, K. L. (1991). The abstraction of form in semantic categories. *Memory & Cognition*, *19*, 1-7.
- SMITH, S. M., & TINDELL, D. R. (1997). Memory blocks in word fragment completion caused by involuntary retrieval of orthographically related primes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *23*, 355-370.

SMITH, S. M., WARD, T. B., & SCHUMACHER, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition*, *21*, 837-845.

TENPENNY, P. L., KERIAZAKOS, M. A., LEW, G. S., & PHELAN, T. P. (1998). In search of inadvertent plagiarism. *American Journal of Psychology*, *111*, 529-559.

WALLACE, W. T., & RUBIN, D. C. (1991). Characteristics and constraints in ballads and their effects on memory. *Discourse Processes*, *14*, 181-202.

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.

WARD, T. B., & SIFONIS, C. (1997). Task demands and generative thinking: What changes and what remains the same? *Journal of Creative Behavior*, *31*, 245-259.

**NOTE**

1. We would like to thank Trisha Tenpenny for suggesting that we conduct this experiment.

**APPENDIX A**  
**Stimuli Used in Consistent Conditions of Experiments 1 and 2**

Category	English Equivalent	Nonword Equivalent	Category	English Equivalent	Nonword Equivalent
Weapons (2 syllables, 7 letters, E ending)			Sports (3 syllables, 8 letters, Y ending)		
	KNIFE	TRENODE		SOCCER	FILLARDY
	GUN	PIRLANE		FOOTBALL	TASSNALY
	BOMB	SPUVONE		BASEBALL	POCOLMEY
	CLUB	GLIPPLE		TENNIS	KARTHELY
	SWORD	KOTSONE		GOLF	SWILINEY
	CANNON	FONTARE		BOXING	RUNBINDY
Insects (1 syllable, 3 letters, R ending)			Four-legged animals (3 syllables, 8 letters, T ending)		
	FLY	BER		HORSE	KEOSLART
	WASP	CIR		COW	MENIDETT
	ANT	DUR		LION	POBOTEET
	BEE	JOR		ELEPHANT	HUMALIRT
	MOSQUITO	NAR		DOG	EKETHANT
	FLEA	WIR		GIRAFFE	SHIWEDAT
Clothing (2 syllables, 5 letters, NG ending)			Fruit (1 syllable, 3 letters, D ending)		
	SHIRT	BEANG		APPLE	ALD
	SOCKS	MOING		PEAR	CED
	PANTS	TAING		BANANA	JID
	COAT	KOENG		GRAPE	LUD
	DRESS	ATING		PLUM	VED
	HAT	HAONG		CHERRY	WUD

**APPENDIX B**  
**Stimuli Used in Experiment 3**

Condition	Pastas	Elements	Analgesics
Natural Endings (I, A) (ON, IUM) (OL, IN)			
	SPAGHETTI	RADON	TYLENOL
	LASAGNA	PLUTONIUM	ANACIN
	FETTUCINI	ARGON	ASPIRIN
	ROTINI	CARBON	BUFFERIN
	PASTINA	RADIUM	PANADOL
	RIGATONI	URANIUM	MIDOL
Mismatched Endings (ON, IUM) (OL, IN) (I, A)			
	TORTIUM	LODOL	PAMPI
	LINGON	GAMIN	ASPIRA
	RONIUM	TIBOL	OUCHI
	PASTON	CHITRIN	CURA
	CATALONIUM	UROL	SISPRRA
	SPAGANON	XENVIN	NARCOMI
Arbitrary Endings (EY, R) (IT, L) (NG, D)			
	UBITINEY	VISIL	KENG
	BIRER	JOTWIT	PLAFID
	CHIR	NATIT	MODED
	ENTAGEY	SUMYL	BESENG
	ASORTLEY	RONKIT	ALYD
	NAPUR	GLIPPEL	WHUN