Journal of Experimental Psychology: Learning, Memory, and Cognition

The Effects of Prior Knowledge on Incidental Category Learning

John P. Clapper

Online First Publication, May 21, 2012. doi: 10.1037/a0028457

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/12/$12.00 DOI: 10.1037/a0028457

The Effects of Prior Knowledge on Incidental Category Learning

John P. Clapper
California State University, San Bernardino

This article describes 5 experiments investigating the role of prior knowledge in incidental category learning. Experiments 1 to 3 showed that prior knowledge improved learning only if the categories in a given set were related to contrasting themes; there was no consistent knowledge effect if the categories were related to the same theme. Experiments 4 and 5 showed that diagnostic verbal labels facilitated the learning of non-thematic categories but provided no additional benefit when the categories were already related to contrasting themes. In terms of the category invention framework proposed by Clapper (2007), these results imply that prior knowledge provides an effective cue for discovering separate categories, as well as helping people segregate the features of those different categories in memory and so improving their recall. The relevance of these results to other types of category learning tasks is discussed.

Keywords: categories, unsupervised learning, incidental learning, prior knowledge

It is widely known that pre-experimental knowledge acquired outside the laboratory can facilitate performance in a variety of category learning tasks. For example, categories that are consistent with prior knowledge are learned more quickly in supervised classification tasks than categories that are inconsistent with such knowledge, and specific features of categories that are consistent with prior knowledge are learned more quickly than features that are neutral or inconsistent with such knowledge (e.g., Heit, 1994; Murphy & Allopenna, 1994; Murphy & Kaplan, 2000; Pazzani, 1991; Rehder & Ross, 2001; Wattenmaker, Dewey, Murphy, & Medin, 1986). Relevant prior knowledge also facilitates learning in unsupervised tasks such as category construction (e.g., Kaplan & Murphy, 1999; Medin, Wattenmaker, & Hampson, 1987; Spalding & Murphy, 1996). Given the ubiquity of such effects, it seems apparent that prior knowledge must play an important role in human categorization.

It also seems reasonable to assume that knowledge might have somewhat different effects on supervised versus unsupervised learning. In supervised tasks, categories are defined in advance by the experimenter; hence, it is natural to focus on how knowledge could facilitate category learning by facilitating the learning of individual features within those categories. In particular, prior knowledge could increase category coherence by specifying meaningful relationships among features, which in turn could make those features easier to learn (e.g., Kaplan & Murphy, 1999; Murphy, 2002; Murphy & Medin, 1985; Wisniewski, 1995). Computational models designed to account for such knowledge effects (e.g., Heit & Bott, 2000; Rehder & Murphy, 2004) provide a kind of formal implementation of this intuition that preexisting associations between knowledge-relevant features, or between those features and an already familiar theme or category, make the features easier to learn.

By contrast, the research reported in this article focuses specifically on the role of prior knowledge in *unsupervised* learning. In unsupervised tasks, people must create their own categories in addition to keeping track of which features are associated with each (Clapper, 2007; Michalski & Stepp, 1983). In principle, prior knowledge could facilitate unsupervised learning not just by helping people to learn which features go with which category but also by helping them to notice or recognize the existence of separate categories in the first place (Clapper, 2007). This "category cuing" role should be particularly important in tasks in which people are initially unaware that categorization is a task goal. In this situation, separate categories could be noticed only by accident or as the result of some external cue, as people would not be searching for them deliberately.

The Exemplar-Memory Task

Investigating these ideas requires a task in which learning is unintentional (incidental) as well as unsupervised. The present experiments employ a variation of the exemplar-memory task first described by Clapper and Bower (2002). In this task, participants are shown a series of training instances, each consisting of a list of verbal descriptors. Their goal is to study each instance and remember its features for an immediate memory test (recognition or cued recall); the next instance is then shown and the cycle is repeated for a specific number of trials. The stimuli in these experiments (e.g., descriptions of fictitious persons) typically fall into two categories based on contrasting clusters of correlated features. Participants are never informed about these categories, but if they happen to discover them on their own they can learn the consistent features of the categories and thereby improve their memory for individual instances. This improvement, relative to a control con-

I thank Xinia Chavez, Melissa Flesher, Jessica Gasca, Charles Griffith, Wade Kidner, Benjamin Malmberg, Claudia Marin, Gabriella Novella, Liz Scholl, Tyler Stickney, and Roseanne Swisher for their assistance in data collection for these experiments.

Correspondence concerning this article should be addressed to John P. Clapper, Department of Psychology, California State University, San Bernardino, 5500 University Parkway, San Bernardino, CA 92407-2397. E-mail: jclappe@csusb.edu

dition in which the training instances lack any consistent category structure, reveals the person's learning.

Clapper (2007) demonstrated that prior knowledge could facilitate unsupervised learning in this task. In particular, when the correlated features that defined two categories were related to familiar "themes" (e.g., suggesting that members of one category were snobby "highbrows" and those of the other were down-to-earth "lowbrows"), people showed much better learning (memory for consistent features) than if the categories did not relate to contrasting themes or stereotypes. These experiments showed better overall learning of categories related to prior knowledge and, within those categories, better learning of features related to prior knowledge than of neutral or unrelated features.

This task provides a convenient way to measure incidental learning without making categorization an explicit goal, while also making it possible to assess the roles that knowledge might play in that overall process. On the other hand, existing computational models of category learning (e.g., Heit & Bott, 2000; Rehder & Murphy, 2004) cannot be applied to this task, as these were intended as models of supervised categorization only. Therefore, a simple qualitative model proposed by Clapper (2007) will be used to provide a framework for the present discussion. Here, I refer to this as the *CBS model* of knowledge effects; the acronym stands for "cuing + binding + segregation" (the rationale for these terms is clarified below). The model begins with an overall conception of the learning process based on the *category invention* theory of Clapper and Bower (1991) and then highlights some obvious ways in which prior knowledge might influence that process.

The Category Invention Theory of Unsupervised Learning

The main assumptions of category invention theory are (a) that people spontaneously attempt to categorize each stimulus they encounter and (b) that they will create a new category whenever they encounter a stimulus that cannot be assigned to any existing category (because it is novel or deviant in some way; e.g., Clapper & Bower, 1991). Once a new category has been created, future stimuli similar to the original "triggering" instance can be assimilated to this category. As this occurs, category norms (Kahneman & Miller, 1986) are acquired so that the person learns which features are consistently present in all (or most) category members and which vary from instance to instance.

Within this framework, creating separate categories is assumed to be the primary way in which people capture the correlational or predictive structure within a given stimulus set. To illustrate, several attributes are perfectly correlated in the stimulus sets used in the present experiments. In principle, one way for a learner to discover this correlational structure would be to store a matrix of all possible interfeature associations in memory and update this matrix over successive training instances. Given sufficient exposure, such "correlation tracking" would eventually reveal which attributes were correlated and which were not within a given stimulus set (Clapper, 2006; Clapper & Bower, 2002). However, one problem with this approach is that the correlation matrix could become very large for a set of complex stimuli that vary along many attributes. An alternative that avoids this type of combinatorial explosion would be to simply divide the stimulus set into explicit subsets (categories), each containing only instances that have a specific cluster of correlated features. Once the stimuli have been partitioned in this way, the learner need only learn a short list of attribute values within each category to capture the predictive structure of the set as a whole (Kaplan & Murphy, 1999).

Thus, given a stimulus set with the appropriate type of correlational structure, inventing separate categories can greatly simplify the task of learning that structure. However, one disadvantage of relying on category invention is that the correlational structure of a given set can remain permanently invisible if the learner assigns the first instance of a potential new category to an already existing category. In this case, the result is a merged or aggregated category that will fit both types of instances; hence, further instances of the potential new category will continue to be assigned to the aggregated category. The lack of a distinct "failure event" means that the learner is trapped in an overaggregated state in which he or she remains persistently blind to the correlational structure of the set. Previous studies have shown strong evidence for this kind of aggregation effect (Clapper, 2006).

The CBS Model of Knowledge Effects

The category invention theory assumes that people capture the correlational structure of a set of stimuli by dividing those stimuli into separate categories and then learning the consistent features within each category. Given this sort of learning process, there are two obvious ways in which prior knowledge of some related theme or stereotype could facilitate the acquisition of new categories (and hence memory for their features).

Category Cuing

As noted, there is no guarantee that people will explicitly search for or notice separate categories in an incidental task. Moreover, even if people are trying to divide a set into categories, the present stimuli vary along many dimensions and so, in principle, could be partitioned in many different ways, only one of which will correspond to the actual correlational structure of that set. Prior knowledge can help learners overcome these problems by cuing them that the stimuli are divisible into separate categories and helping them determine to which category each stimulus belongs. For example, if a participant is studying descriptions of fictitious persons, some of whom are described as liking the opera and others as preferring professional wrestling, this might call to mind preexisting stereotypes of highbrows versus lowbrows and lead the learner to hypothesize that the descriptions being studied fall into two categories based on this distinction.

Although thematic knowledge may provide the cue that triggers new categories, the actual effects of such cues should be independent of thematic relevance. Indeed, relatedness to prior knowledge is only one type of possible category cue; others include diagnostic verbal labels, instructional or contextual cues, and a variety of other devices. But regardless of how people come to divide a stimulus set in a particular way, the effect should be the same—namely, it becomes possible to capture the correlational structure of that set simply by learning a short list of consistent features associated with each category. This benefit should apply equally to all consistent features within a category, not just those that happen to be related to prior knowledge. I would expect essentially the same effect from adding a diagnostic verbal category label to each instance, a prediction explored in some of the experiments reported below.

Feature Facilitation

In addition to the category cuing effect just discussed, prior knowledge could also provide a benefit by facilitating the learning of individual features within each category. In contrast to category cuing, such feature facilitation effects should be larger for features that are relevant to prior knowledge than for irrelevant or neutral features. For example, if we know that a man is wealthy it might help us remember what kind of car he drives, but that knowledge would do little to help us remember his hair color or the month in which he was born. The CBS model assumes that there are two ways in which feature facilitation effects might occur (Clapper, 2007).

Feature binding. It has often been proposed (e.g., Kaplan & Murphy, 1999; Murphy & Medin, 1985) that prior knowledge specifies various types of interfeature relations within a category. This in turn might make those features easier to remember due to the sort of mutual cuing or coactivation effects often reported in the memory literature (e.g., Roediger & McDermott, 1995). A key point about such *feature binding* effects (Clapper, 2007) is that improved memory for the features of a given category is a direct outcome of factors internal to that category itself (e.g., its relation to prior knowledge, preexisting associations among its features), independent of whatever other categories the person is learning or already happens to know.

Feature segregation. A second reason why relevant features of some category (A) might be learned better in a prior knowledge condition is due to the effects of knowledge, not on A itself but on the other category being learned (B)—or more precisely, on the relationship between A and B. To illustrate, consider a task in which participants are studying descriptions of fictitious persons that fall into categories corresponding to highbrows versus lowbrows (as in the examples used earlier). In this situation, there are two reasons why people might find it easy to learn that that the highbrows' favorite snack is, say, caviar rather than pizza. The first is that this preference is consistent with what they already know about highbrows, and this preexisting connection makes the new association easier to acquire—a feature binding effect. However, a second possibility is that knowledge might help not so much by strengthening links to correct features as by inhibiting links to incorrect features. If one is learning two categories of person descriptions, and each has a different default value for the "preferred snack" attribute, there will naturally be some possibility for confusion and ordinary associative interference during the learning process (e.g., J. R. Anderson & Bower, 1973; Keppel, 1968; Postman, 1971). Being able to relate the two categories to opposing themes or stereotypes (highbrows vs. lowbrows) could reduce interference and help people to avoid mixing up their features; in the highbrow/lowbrow case, for example, it should be easy to avoid falsely recalling that John likes pizza rather than caviar if we know that John is a highbrow and regards pizza as a lowbrow food. Clapper (2007) referred to this form of feature facilitation as a feature segregation effect, because in this case knowledge improves memory for A by reducing interference from B (i.e., by making it easier to segregate A features from B features in memory).

Combined model. Given these different types of potential knowledge effects, the overall model can be abbreviated as "knowledge effect = category cuing effect + feature facilitation effect," where it is understood that feature facilitation should be dependent on category cuing (i.e., people can only associate features with categories if they first recognize the existence of those categories). The feature facilitation effect itself is a combination of feature binding and feature segregation, as described above, so the complete CBS model can be summarized as "knowledge effect = category cuing effect + feature binding effect + feature segregation effect."

This formulation is intended to convey the idea that the component knowledge effects can be distinguished in principle and that all are expected to make a positive contribution to overall learning. It is not intended as a claim that the component effects are strictly additive or that they are independent in a statistical sense. As noted, for example, feature binding and segregation effects are expected to occur only if people have already invented separate categories; thus, they are not expected to occur independently of category cuing effects.

Relation to Prior Exemplar-Memory Results

The pattern of knowledge effects found by Clapper (2007) is readily interpreted within this framework. First, people learned the consistent features of categories in thematic conditions better than those in comparable non-thematic conditions, even if those features were not directly related to the themes. This is consistent with a category cuing effect, which should benefit all consistent features equally. Second, people in the thematic condition learned thematically relevant features better than neutral or irrelevant features. This is consistent with a feature facilitation effect, which should favor thematically relevant over irrelevant features in ease of learning.

More detailed results suggested that the feature facilitation effects in this task were primarily segregation effects, with little evidence for enhanced binding. In one experiment, the features of a neutral or non-thematic category, A, were learned better if the other category, B, was thematic than if B was neutral; moreover, the features of A corresponding to thematic features of B were remembered better than other consistent features of A. In other words, this experiment demonstrated a knowledge effect (of B) on a non-thematic category (A) and specifically on neutral features of A corresponding to thematic features of B. This cannot be a binding effect as defined above, as that refers to an effect of prior knowledge about a given category on memory for the features of that category itself (i.e., knowledge related to B improving memory for features of B). Here, by contrast, there is an effect of prior knowledge of B on learning the features of A; within the current framework, that could only be a segregation effect. Additionally, in the condition in which B was thematic and A was neutral, the features of B were not remembered any better than those of A; thus, there was no extra benefit for actually being the thematic category as opposed to simply being easy to tell apart from one, as would be expected given a significant binding effect. All that seemed to matter in this experiment was that having thematic knowledge about one of the categories made them both easy to distinguish overall as well as at the level of individual features.

Why No Binding Effects?

So far, I have defined a binding effect as the direct facilitation of memory for a feature due to thematic knowledge related to the category to which that feature belongs. Thus, a binding effect would be expected to facilitate performance by increasing the strength or availability of the correct response (feature) from memory, given the category as a retrieval cue. It seems reasonable to assume that this would be particularly advantageous when potential responses have low baseline availability, so that increasing the activation of the correct response would lead directly to an improvement in memory. On the other hand, if both correct and incorrect responses were already highly available from memory (at ceiling activation), response generation would no longer be an issue and response discrimination (segregation) would become the major factor determining performance.

In the task used in Clapper (2007), people were provided with both correct and incorrect responses as alternatives on the forced-choice recognition tests. Hence, inability to generate potential responses was not a limiting factor in this experiment. The difficult part, presumably, was selecting the correct response for the current instance. In this situation, thematic knowledge would mainly be useful for excluding or eliminating incorrect alternatives. Thus, if we are trying to remember the last person's favorite food, and we know he or she was a highbrow, we can eliminate pizza from consideration and select caviar by default (assuming those are the only choices). Importantly, it should not matter whether the knowledge in question actually relates to the category whose features the person is currently trying to recall. Thus, if we have two categories of "average" people, one of whom likes pizza and the other pork, prior knowledge does not help us select a correct response. But if we compare that non-thematic condition to a condition in which one of the categories has been altered so that it relates to a familiar theme (e.g., one of the non-thematic categories is retained while the other is changed to wealthy highbrows), we will find that this knowledge helps the participant recall both the features of the thematic category itself and the features of the other, non-thematic category (Clapper, 2007). Thus, when recalling the thematic category they can use the fact that the person was a highbrow to eliminate pizza as a possible response, and when recalling the non-thematic category they can use the fact the person was not a highbrow to eliminate caviar as a possible response.

One could argue that the forced-choice recognition task used in Clapper (2007) was particularly unfavorable to binding effects due to that fact that all alternatives were explicitly presented during each test, reducing the relevance of any potential generation effect. The version of the exemplar-memory task used in the present experiments employed cued recall instead of recognition as an index of memory. This task required people to generate potential responses for themselves rather than merely select among pre-presented alternatives, which could increase the relevance of possible binding effects in the present task.

Overview of the Present Experiments

The present experiments attempted to extend these earlier results by identifying and measuring the contributions of each of the three component knowledge effects described in the model above. The first three experiments compared conditions in which people learned categories related to different themes, as in the studies discussed so far, to conditions in which both categories were related to the same theme. The CBS models expects no overall knowledge effects in the same-theme condition, due to a lack of category cuing effects. The model assumes that thematic knowledge can only improve recall if people are able to create separate categories, and having both categories related to the same theme provides no obvious category cue to help them do that. Thus, a lack of knowledge effects in the same-theme conditions, combined with strong and reliable knowledge effects in the opposite-theme conditions, would provide support for the model in general and for its assumptions about category cuing in particular. Experiment 4 sought further evidence for the distinction between category cuing and feature facilitation by comparing the effects of prior knowledge to that of a "pure" category cuing manipulation, namely, verbal labels that specify the category membership of each training instance and so "force" people to notice separate categories. The CBS model predicts that the total knowledge effect should be larger than the pure category cuing effect (because the latter is only one component of the former) and that the two effects should be non-additive when combined (i.e., adding diagnostic labels to the training instances in a thematic condition should provide no additional improvement in learning). Finally, Experiment 5 sought specific evidence for feature binding by testing for a knowledge effect in a same-theme condition in which separate categories were provided a priori by diagnostic verbal labels. Given that people have already created separate categories in response to these labels, the question is whether prior knowledge would then facilitate thematic feature memory in the same-theme condition as it does in the opposite-theme condition. If so, this would provide strong evidence for a feature binding effect in the same-theme condition; on the other hand, the lack of such an effect would suggest that segregation effects loom much larger than binding effects in the present task.

Experiment 1

This experiment compared learning in three conditions. In the same-theme condition, people saw examples of two categories related to the same theme but with different consistent features (e.g., two categories of highbrows, one of which prefers caviar and the other escargot). In the opposite-theme condition, the categories were related to contrasting themes (e.g., highbrows vs. lowbrows). A neutral or no-theme condition was included as a baseline. As noted above, a knowledge-effects model based on category invention does not expect significant benefits in the same-theme condition; relating different categories to the same theme provides no obvious category cuing effect, and so there is no reason to expect performance in the same-theme condition to exceed that in the no-theme condition. On the other hand, if knowledge effects in this task do not depend on inventing separate categories but only on, say, the existence of prior associations among thematic features, people should show better learning in both thematic conditions compared to the no-theme condition.

Method

Participants. A total of 58 undergraduate students of California State University, San Bernardino, participated in exchange for extra credit in selected psychology classes.

Procedure. Participants were tested in a group computer lab seating up to 20 people. The experiment consisted of a total of 48 trials. Each trial consisted of a study phase followed by a test phase. During the study phase, a training instance consisting of a 12-item verbal feature list describing a fictitious person was presented in the middle of the computer screen (see Figure 1a). The label at the top of the list consisted of the person's initials printed in capital letters; the 12 feature descriptors appeared below. Each feature descriptor consisted of a short phase identifying a specific feature (value of a particular attribute dimension) of that person (e.g., *favorite activity: sailing*, *lives in: Tucson*). The attributes were shown in the same order on all trials (this order was randomized across participants). Each training instance was presented as a description of a different fictitious person. A given training instance remained on the screen for 18 seconds.

The test phase of each trial consisted of a series of four cued-recall tests for the attribute values of the training instance shown on that trial (see Figure 1b). The particular dimensions tested on a given trial were selected randomly from among the 12 attribute dimensions, with the constraint that three of the dimensions tested on each trial were consistent dimensions and one was a variable dimension (see below). For each test, the label (initials) of the current training instance (person) was shown at the top of the screen; below it was shown the name of a particular attribute dimension. The participant attempted to type in the correct value of that attribute (i.e., the value possessed by the current instance). After this response had been entered, the correct response was displayed beneath (see Figure 1c). The same procedure was followed for all four tests, after which the next training instance appeared and a new study-test trial began.

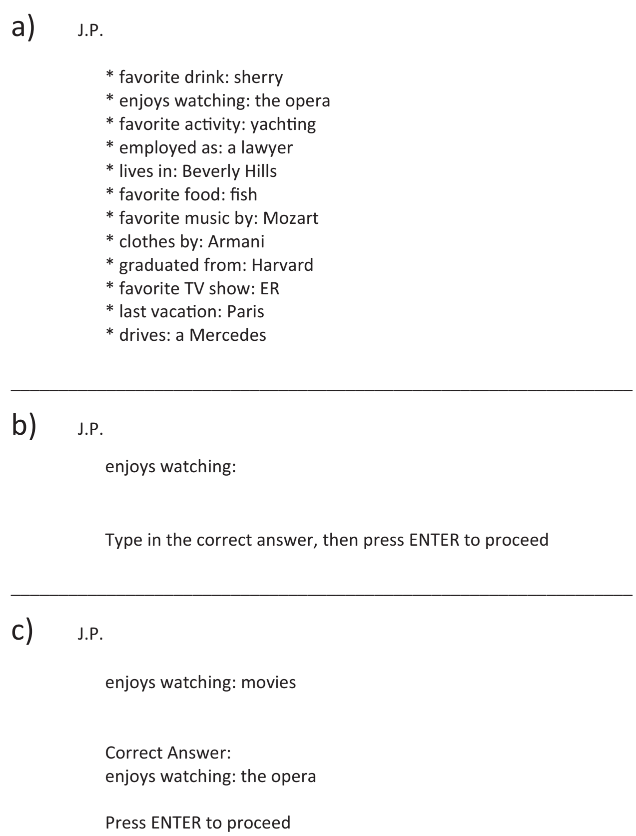

Figure 1. Stimulus displays for the exemplar-memory task. Panel a shows a training instance presented during the study phase of a typical trial. Panels b and c show a sample recall test presented during the test phase of a typical trial.

Materials. Each training instance was composed of 12 attribute dimensions, each of which had seven possible values. The attribute dimensions referred to specific characteristics, such as the person's interests, education, hobbies, and tastes (see the Appendix). Of the seven values of each dimension, three were thematic values (these are the first three listed for each dimension in the Appendix). The first two were related to one theme and the third value was related to the opposite theme, for example, "drinks fine wine," "drinks sherry," and "drinks beer" (the first two of these relate to a highbrow theme while the third relates to a contrasting lowbrow theme; see below). The remaining four values of each dimension were all neutral values that were not obviously related to either theme, for example, "drinks coffee," "drinks soda."

To guard again the idiosyncrasies of particular stimulus sets, three different sets were used (one for any given participant; see the Appendix). These stimulus sets differed in the particular attribute dimensions used to describe the training instances and in the nature of the contrasting themes (highbrow/lowbrow; male/female; senior/youth). The theme assigned to have two values for each dimension was the one judged to be more marked or salient (i.e., that would be easiest to recognize or stereotype even in the absence of its opposite), highbrows for the highbrow/lowbrow set, males for the male/female set, and seniors for the senior/adolescent set. The third thematic value was related to the opposite theme within each set (lowbrows, females, or youth). The themes were never mentioned explicitly in the instructions; individual participants were left to notice these themes (or not) for themselves.

Design. Participants were randomly assigned to one of three conditions, referred to as the opposite-theme ($N = 21$), no-theme ($N = 19$), and same-theme ($N = 18$) conditions. In each condition, nine of the 12 stimulus dimensions had perfectly correlated values and the remaining three dimensions were completely uncorrelated. The correlated dimensions defined two categories within a given stimulus set; these dimensions had the same values across all instances of a given category, while the uncorrelated dimensions varied independently across instances. The two categories in the same-theme condition had correlated values that were related to the same theme. For example, assuming that values 1 and 2 are both related to the highbrow theme, the categories could be represented as Category A = 111111111XXX and Category B = 222222222XXX, where the Xs in the last three positions indicate uncorrelated dimensions, each of which could take on four possible neutral values (values 4–7 in the Appendix) within a given category. The opposite-theme conditions had correlated values that related to contrasting themes. If value 1 refers to the highbrow theme and value 3 refers to the lowbrow theme, the categories could be represented as Category A = 111111111XXX and Category B = 333333333XXX. The categories in the no-theme conditions bore no obvious relation to the themes; these can be denoted as Category A = 444444444XXX and Category B = 555555555XXX, where values 4 and 5 are neutral with respect to the themes and the Xs have the same interpretation as described above.

A total of 48 instances was presented over the course of the experiment, 24 from each category. The categories were shown in pseudo-random alternation, such that no more than three examples of the same categories could be shown in a row. The same abstract sequence was used for all participants within a given condition, but due to the counterbalancing of the materials (see below) each participant saw different stimuli.

Counterbalancing. As described above, three different stimulus sets were used to ensure the generality of the results. Each set was shown to an equal number of participants and was based on a different pair of contrasting themes (senior/youth, male/female, and lowbrow/highbrow). Particular stimulus attributes were assigned randomly to abstract roles (e.g., correlated vs. uncorrelated) in the experimental design; this was done separately for each participant. The order of attributes in the list display was also randomized separately for each participant.

Results and Discussion

The data from this experiment consist of the immediate recall of the consistent and variable features of each training instance. Except where explicitly noted, the bottom 25% of participants were excluded from each condition in all analyses reported below as well as in Figure 2 and Tables 1 and 2. This is an unbiased procedure and did not change the overall pattern of results in any case, but it did tend to slightly increase the power of the statistical tests. (The same procedure is followed for all of the experiments reported in this article.) The most important data are the consistent feature recalls, displayed in Figure 2. The condition means and planned comparisons are shown in Tables 1 and 2. In each case, statistics were computed over the second half of trials to assess asymptotic performance.

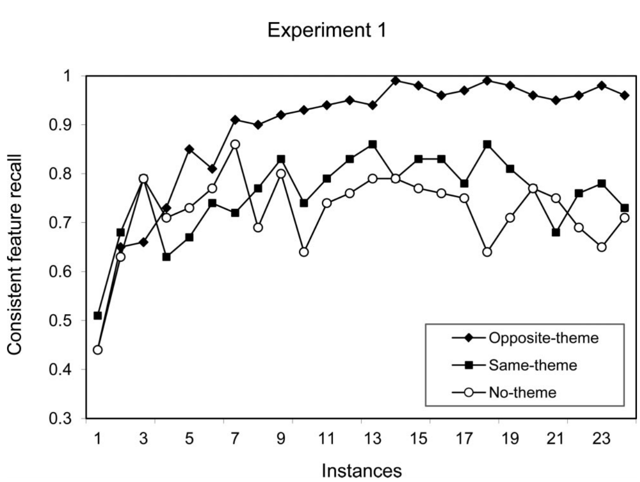

Figure 2. Consistent feature recall data from Experiment 1, plotted over trials. The top line shows data from the opposite-theme condition, the middle line shows data from the same-theme condition, and the bottom line shows data from the no-theme condition.

Table 1
Recall Memory Data From Experiment 1

Condition	Dependent variable	*M*
Opposite-theme	Consistent feature recall	.96
	Variable feature recall	.80
Same-theme	Consistent feature recall	.79
	Variable feature recall	.66
No-theme	Consistent feature recall	.73
	Variable feature recall	.63

Note. Condition means in all tables are computed over the second half of trials only.

One-way analyses of variance (ANOVAs) revealed significant differences between conditions for both consistent, $F(2, 55) = 22.74$, $p = .000$, and variable dimensions, $F(2, 55) = 7.30$, $p = .002$. Planned comparisons showed significantly greater recall of both consistent and variable features in the opposite-theme condition than in the same-theme and no-theme conditions (see Table 2). The same-theme and no-theme conditions did not differ in variable feature recall, but the difference in recall of consistent features was marginally significant ($p = .052$; see Table 2). This difference did reach statistical significance according to an independent-samples Mann–Whitney U test ($p = .032$).

In summary, the data provide preliminary evidence for knowledge effects in both opposite- and same-theme conditions, as both showed greater recall of consistent features compared to the no-theme condition. The finding of knowledge effects in the same-theme condition was not predicted by the CBS model, as being related to the same theme provides no obvious cue to separate categories. However, it is clear that the knowledge effects in the present study were much larger in the opposite-theme condition than in the same-theme condition. Given this difference, it is possible that the smaller effect in the same-theme condition had a different cause than that observed in the opposite-theme condition, one that did not depend on inventing separate categories. Thus, learners in the same-theme condition might have noticed that all the people they studied appeared wealthy in most of their attributes, which might in turn have made it easier to recall those attributes even without inventing separate categories. For example, it might be easier to recall what kind of car someone has merely by knowing that it is an expensive car; such knowledge narrows the list of candidates and may also facilitate response generation.

Experiment 2

According to the CBS model, the benefits observed in the opposite-theme conditions of Experiment 1 should reflect some mix of category cuing and feature facilitation effects. One difficulty in attempting to disentangle these two effects stems from the fact that all the consistent features of the thematic categories in Experiment 1 were related to the target themes. Because these features were all thematically relevant, it is impossible to tell whether better recall was due to enhanced category cuing, feature facilitation, or some combination of the two. In order to distinguish these different possibilities, a situation in which the model predicts both effects (category cuing and feature facilitation) must be compared to a situation in which it predicts only one (category

Table 2
Key Statistical Comparisons From Experiment 1

Dependent variable	Comparison	*t*	*df*	*SE*	*p*
Consistent feature recall	Opposite- vs. same-theme	8.87	27	.020	.000
	Opposite- vs. no-theme	9.94	28	.023	.000
	Same- vs. no-theme	2.04	25	.028	.052
Variable feature recall	Opposite- vs. same-theme	2.91	27	.020	.007
	Opposite- vs. no-theme	3.36	28	.047	.001
	Same- vs. no-theme	0.63	25	.041	.534

cuing). One goal of Experiments 2 and 3 was to enable such a comparison by assessing knowledge effects separately for thematic versus neutral features within a given category.

Previous experiments have provided evidence that knowledge facilitates learning of both thematic and neutral features of relevant categories. Kaplan and Murphy (1999), for example, showed significant knowledge effects for both types of features in a different unsupervised task. In the present task, Clapper (2007) demonstrated significant effects for thematic and neutral features of thematically relevant categories, showing in addition that the effect was larger for thematic than for neutral features. However, the knowledge effect for neutral features failed to reach statistical significance in a second experiment reported in that article and so stands in need of replication. The method used in those experiments was also slightly different than that employed here, as feature memory was tested with forced-choice recognition rather than cued recall.

Comparing knowledge effects for thematic versus neutral consistent features is also relevant to determining whether the small improvement in recall in the same-theme condition of Experiment 1 was due to knowledge somehow cuing people to learn separate categories or to some overall memory effect, as suggested above. In particular, improved recognition of separate categories should lead to improved recall, not only of thematic features but also of neutral consistent features. At the same time, it is worth noting that making some of the consistent features neutral rather than thematically relevant might reduce the scope for guessing strategies, decreasing any advantage for the same-theme relative to the no-theme condition.

Experiment 2 assessed knowledge effects on thematic versus neutral features in two conditions similar to the opposite-theme and no-theme conditions from Experiment 1. Assuming that both neutral and thematic features should benefit from category cuing, greater recall of both in the thematic (opposite-theme) condition would suggest better recognition of separate categories in that condition. Greater recall of thematic compared to neutral features would then provide evidence for an additional feature facilitation effect in that condition. Experiment 3 extended these comparisons by adding a same-theme condition, as described above.

Method

Participants. Thirty-five undergraduate students of California State University, San Bernardino, participated in exchange for extra credit in selected psychology classes.

Procedure. The procedure was similar to that of the Experiment 1. The main difference was that the amount of time each study list was shown decreased over trials in the present experiment. The lists were displayed for 18 seconds per trial for the first 24 trials, after which the exposure time decreased by 0.5 s/trial, declining to only 6 seconds by the 48th and final trial. This reduction in study times was included to increase the difficulty of later trials and, hopefully, to increase observed differences between conditions.

Materials and design. The materials were similar to those of Experiment 1 (see the Appendix). Participants were randomly assigned to two conditions referred to as the thematic and non-thematic conditions; these were similar to the opposite-theme and no-theme conditions from Experiment 1. Using the same notation as above, the category structures in the thematic condition may be denoted as A = 111111444XXX and B = 333333555YYY, where 1s and 3s are thematic values, 4s and 5s are neutral values, and the Xs or Ys indicate uncorrelated dimensions with three possible values (1, 4, or 5 for X and 3, 5, or 6 for Y). The category structures in the non-thematic condition may be denoted as A = 444444444XXX and B = 555555555YYY.

As in Experiment 1, the stimuli were presented in a pseudo-random sequence (no more than three instances of the same category in a row). The same counterbalancing procedures were also used in the present experiment (e.g., three different stimulus sets, random assignment of stimulus features to roles in the experimental design).

Results and Discussion

Consistent feature recall data for this experiment are shown plotted over trials in Figure 3, and means and key tests of statistical significance are summarized in Tables 3 and 4. As shown in these tables, both consistent and variable features were recalled significantly better in the thematic condition than in the non-thematic condition. Both thematic and neutral consistent features in the thematic condition were recalled better than the corresponding features in the non-thematic condition. Thematic consistent features were recalled better than neutral consistent features within the thematic condition, and neutral consistent features were recalled better than variable features. Within the non-thematic condition, consistent features were recalled better than variable features.

Above, it was argued that recognizing separate categories should improve recall of all the consistent features of those categories, including consistent features unrelated to the themes that may have triggered the categories in the first place. The present demonstration of a knowledge effect on neutral features provides evidence for such a category cuing effect and reinforces similar results reported by Kaplan and Murphy (1999) and Clapper (2007).

The other major result was that thematic features showed a significant advantage over neutral features within the thematic condition itself. As noted above, this provides evidence that prior knowledge helps people learn the individual features of new categories. Thematically relevant features should be the primary beneficiaries of such feature facilitation effects; hence, thematic consistent features should be learned somewhat more easily than neutral consistent features, as in the present data.

Experiment 3

The goal in the next experiment was to replicate Experiments 1 and 2 within a single combined study. The design was similar to Experiment 1, except that the two thematic conditions included both thematic and neutral consistent features, as in Experiment 2. Experiment 1 demonstrated a strong effect of prior knowledge when the categories were related to contrasting themes and a smaller effect when both were related to the same theme. It was unclear whether the latter result was due to people somehow using thematic knowledge to help them learn individual categories or whether some other factor, such the more restricted range of possible attribute values in the same-theme condition, led to the slight improvement observed there. Having people learn categories containing both thematic and neutral consistent features should make it possible to distinguish between these two possibilities by determining (a) whether the same-theme benefit extends to situations in which a smaller proportion of a category's consistent features is thematically relevant and (b) assuming that any benefit is observed in the same-theme condition, whether that benefit occurs for both neutral and thematic consistent features.

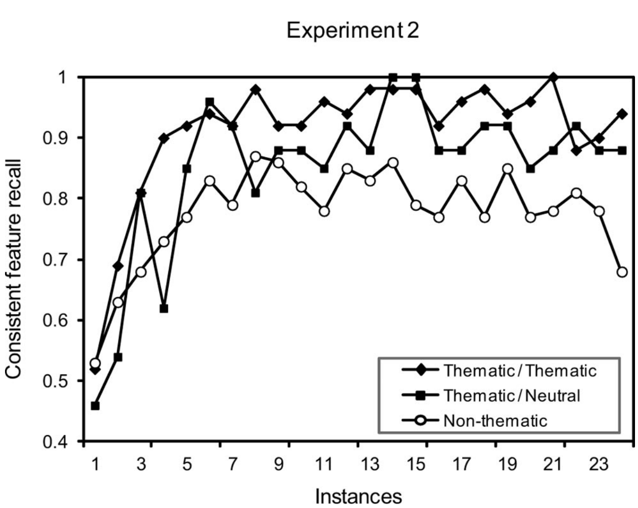

Figure 3. Consistent feature recall data from Experiment 2, plotted over trials. The top line shows data from thematic features within the thematic condition, the middle line shows data from neutral features within the thematic condition, and the bottom line shows data averaged over all consistent features in the non-thematic condition.

Table 3
Recall Memory Data From Experiment 2

Condition	Dependent variable	*M*
Thematic	Consistent feature recall (thematic)	.95
	Consistent feature recall (neutral)	.91
	Variable feature recall	.77
Non-thematic	Consistent feature recall (all)	.79
	Variable feature recall	.65

Method

Participants. A total of 45 undergraduate students of California State University, San Bernardino, participated in exchange for extra credit in selected psychology classes.

Procedure. The procedure was similar to that of the previous experiments. Note that the study time reduction over trials from Experiment 2 was retained in this experiment; lists were displayed for 18 s/trial for the first 24 trials, after which exposure time decreased by 0.5 s/trial, reaching a minimum of 6 seconds on the final (48th) trial.

Materials and design. The materials and design were essentially identical to those of Experiment 1, except that the categories in the two thematic conditions had both thematic and neutral consistent values, as in Experiment 2. The abstract structure of these stimuli may be represented as follows: in the same-theme condition, Category A = 111111444XXX and Category B = 222222555XXX; in the opposite-theme condition, Category A = 111111444XXX and Category B = 333333555XXX; and in the no-theme condition, Category A = 444444444XXX and Category B = 555555555XXX. Here, values 1 and 2 denote different thematic values pertaining to the same theme (highbrow, male, or senior), value 3 is a thematic value pertaining to the opposite theme (lowbrow, female, or adolescent), and values 4 and 5 are neutral values not obviously related to either theme. The Xs in the final three positions refer to the four neutral values (4, 5, 6, or 7). All counterbalancing procedures were the same as in the previous experiments.

Results and Discussion

Consistent feature recalls for the opposite-theme condition versus the other two conditions (pooled) are shown plotted over trials in Figure 4;[1] data for thematic versus neutral dimensions are shown separately in this figure. Condition means and key statistical tests are shown in Tables 5 and 6. One-way ANOVAs revealed significant differences in asymptotic recall averaged over the second half of trials for thematic, $F(2, 30) = 5.00$, $p = .013$, and variable dimensions, $F(2, 30) = 3.90$, $p = .031$; the difference for neutral dimensions did not reach conventional levels of statistical significance, $F(2, 30) = 2.19$, $p = .129$. More detailed planned comparisons (see Table 6) showed that these effects were caused by the opposite-theme condition enjoying significantly better recall than the other two conditions, which did not differ between themselves. Inspection of Tables 5 and 6 shows that average

[1] Data for two participants were discarded because they did not complete the experiment.

performance was very similar in the latter two conditions (hence they are combined in Figure 4), whereas the opposite-theme condition showed significantly greater performance for thematic and variable dimensions. The comparison for neutral dimensions was also significant when the opposite-theme condition was compared to the pooled data from the other two conditions.

In summary, this experiment provided further evidence for the benefit of prior knowledge in the opposite-theme condition compared to the no-theme condition. The pooled comparison just described also replicates the neutral-feature advantage for thematic categories found in Experiment 2. On the other hand, the weak knowledge effect in the same-theme condition of Experiment 1 did not recur in the present experiment. This was not due to the data trimming procedure used in these experiments, as including all participants' data did not change the results. When thematic feature data from the same-theme and no-theme conditions of the present experiment were pooled with the corresponding data from Experiment 1, the combined analysis showed no evidence of any knowledge effect in the same-theme condition, $t(51) = 0.63$, $p = .530$. Taken together, the two experiments provide little support for a reliable effect of prior knowledge in the same-theme condition, especially when contrasted with the robust knowledge effects consistently found in opposite-theme conditions here and in Clapper (2007). Another difference from Experiment 1 is that the difference between thematic and neutral dimensions failed to reach significance in the opposite-theme condition of the present experiment, $t(10) = 0.50$, despite a small difference in means in the expected direction. The reason for this non-result is unclear, particularly as significant differences between thematic and neutral feature were found in both Clapper (2007) and the present Experiment 2.

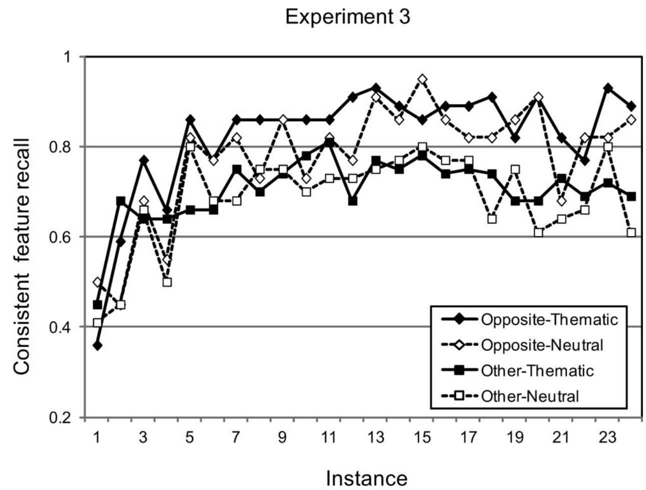

Figure 4. Consistent feature recall data from Experiment 3, plotted over trials. The upper solid line shows data from thematic features within the opposite-theme condition; the upper dotted line shows data from the neutral features within the opposite-theme condition. The lower solid line shows data from thematic features of the same-theme condition pooled with the corresponding features from the no-theme condition. The lower dotted line shows data from neutral features of the same-theme condition pooled with the corresponding features from the no-theme condition.

Experiment 4

In the first three experiments, people in opposite-theme conditions showed consistent knowledge-based facilitation for both thematic and neutral features (generally more for the former than the latter but significant for both). By contrast, people in the same-theme conditions did not show consistent knowledge effects across experiments. The significant improvement observed in the opposite-theme conditions for all consistent features suggests that the themes were serving as category cues in that condition. The fact that this benefit was generally larger for thematic than for neutral features further suggests that the themes were providing additional facilitation in learning individual features. In other words, the data so far are consistent with the broad distinction between category cuing and feature facilitation posited by the CBS model, and they suggest that both may operate when the categories to be learned are related to contrasting themes.

The next experiment aimed to provide further evidence for this distinction and, more specifically, for the model's prediction that that the total impact of prior knowledge on the recall of any particular feature would reflect a combination of these two effects (i.e., total knowledge effect = category cuing effect + feature facilitation effect). Note that this formulation makes a simple and testable prediction, namely, that the total knowledge effect should be greater than the category cuing effect alone, assuming a non-zero feature facilitation effect. This implies that "pure" manipulations of category cuing should improve learning but the potential impact of such manipulations

Table 4
Key Statistical Comparisons From Experiment 2

Comparison	*t*	*df*	*SE*	*p*
Between-groups				
Thematic dimensions	5.78	24	.03	.000
Neutral dimensions	3.45	24	.04	.002
Variable dimensions	2.29	24	.05	.031
Within-groups comparisons (thematic condition)				
Thematic vs. neutral	1.97	12	.02	.073
Thematic vs. neutral (all participants)	2.97	16	.03	.009
Thematic vs. variable	5.91	12	.03	.000
Neutral vs. variable	3.51	12	.04	.004
Within-groups comparisons (non-thematic condition)				
Consistent vs. variable	6.86	12	.02	.000

Table 5
Recall Memory Data From Experiment 3

Condition	Dependent variable	*M*
Opposite-theme	Consistent feature recall (thematic)	.88
	Consistent feature recall (neutral)	.85
	Variable feature recall	.67
Same-theme	Consistent feature recall (thematic)	.72
	Consistent feature recall (neutral)	.73
	Variable feature recall	.47
No-theme	Consistent feature recall	.72
	Variable feature recall	.50

should be less than that of thematic relatedness, because category cuing is only one component of the overall thematic effect. For example, adding diagnostic verbal labels that reveal the category membership of each instance has been shown to improve learning in a variety of supervised and unsupervised tasks, including the present exemplar-memory task (Clapper & Bower, 2002; Yamauchi, 2005). In the present context, it would seem reasonable to assume that such labeling effects represent a relatively pure manipulation of category cuing. Given this assumption, the CBS model predicts that label manipulations should have significantly weaker effects on learning than otherwise equivalent manipulations of thematic knowledge.

The present experiment attempted to distinguish category cuing from feature facilitation effects by comparing the effects of thematic relatedness to that of diagnostic labels. Two predictions follow from the discussion so far. First, all else being equal, the effects of thematic relevance should be greater than the effects of diagnostic verbal labels, because the former involves both category cuing and feature facilitation and the latter involves only category cuing. (Importantly, this should be true only for thematic features of thematic categories; neutral features of thematic categories should benefit from category cuing but not from feature facilitation, so recall of these features should be the same as in a diagnostic-labels-only condition.) Second, whereas adding diagnostic labels to non-thematic categories should improve learning, adding such labels to thematic categories should have no significant effect, as any category cuing benefit provided by the labels would be redundant with that already provided by the contrasting themes.

Method

Participants. A total of 80 undergraduate students of California State University, San Bernardino, participated in exchange for extra credit in selected psychology classes.

Procedure, materials, and design. The procedure was similar to that of the previous experiments. Lists were displayed for 18 s/trial for the first 24 trials, after which exposure time decreased by 0.5 s/trial, reaching a minimum of 6 seconds on the final (48th) trial. The main procedural difference from earlier experiments was that participants were asked to recall the instance label on each trial. The label test always occurred first, prior to the four feature tests.

Two main independent variables were manipulated in a between-groups factorial design. First, half the participants saw categories based on familiar themes and the other half saw non-thematic categories: Category A = 111111444XXX and Category B = 333333555XXX in the thematic conditions; Category A = 444444444XXX and Category B = 555555555XXX in the non-thematic conditions. (As previously, values 1 and 3 denote thematic values pertaining to opposite themes and values 4 and 5 are neutral values not obviously related to either theme.) The second independent variable concerns the verbal labels. Half the participants saw training stimuli in which each instance received a different verbal label (initials shown at the top of each training instance as in the previous experiments). The other half saw training stimuli in which each instance was given one of two diagnostic labels (“GROUP 1” or “GROUP 2”) depending on their category membership. Crossing these two factors resulted in four groups, henceforth referred to as the thematic/diagnostic, thematic/non-diagnostic, non-thematic/diagnostic, and non-thematic/non-diagnostic conditions.

As in previous experiments, training instances were shown in the same order to all participants within a given condition. Because features were randomly assigned to roles in the abstract design for each participant, different actual training instances were shown to

Table 6
Between-Groups Planned Comparisons From Experiment 3

Type of dimension/condition	*t*	*df*	*SE*	*p*
Thematic dimensions				
Opposite- vs. same-theme	2.48	18	.061	.023
Opposite- vs. no-theme	3.37	22	.043	.003
Same- vs. no-theme	0.29	20	.060	.940
Opposite-theme vs. others (pooled)	3.21	31	.046	.003
Neutral dimensions				
Opposite- vs. same-theme	1.55	18	.075	.139
Opposite- vs. no-theme	1.99	18	.074	.061
Same- vs. no-theme	0.37	20	.078	.709
Opposite- theme vs. others (pooled)	2.09	31	.064	.045
Variable dimensions				
Opposite- vs. same-theme	2.43	18	.082	.026
Opposite- vs. no-theme	2.55	22	.068	.018
Same- vs. no-theme	0.30	20	.083	.769
Opposite-theme vs. others (pooled)	2.81	31	.065	.008

each participant. All other counterbalancing procedures were the same as in previous experiments.

Results

The data consisted of recall performance for thematic, neutral, and variable dimensions, as well as the instance labels, within each experimental condition. Asymptotic recall performance (averaged over the second half of trials) for thematic and neutral dimensions across the four conditions is shown in Figure 5. All the analyses below are based on asymptotic recall. Condition means are shown in Table 7.

The main predictions for this experiment concerned the effects of thematic knowledge and diagnostic labels on the recall of thematic and neutral consistent features. Beginning with thematic dimensions, a 2 × 2 ANOVA revealed a significant main effects of theme, $F(1, 56) = 26.71$, $p = 000$, and a marginally significant effect of label, $F(1, 56) = 3.48$, $p = .067$. The interaction was also significant, $F(1, 56) = 4.62$, $p = .036$. A similar pattern was shown by the neutral dimensions. Both the main effect of theme and the interaction term were statistically significant, $F(1, 56) = 11.073$, $p = .002$ and $F(1, 56) = 6.48$, $p = .014$, respectively. However, the main effect of labels failed to reach significance for neutral dimensions, $F(1, 56) = 1.33$, $p = .253$.

Because the interaction terms were significant in both analyses, planned comparisons were carried out to test for differences between individual conditions. There was no effect of category label on the thematic conditions for either thematic or neutral dimensions, $t(28) = 0.25$, $p = .805$ and $t(28) = 1.07$, $p = .295$. Label diagnosticity did have a significant effect on the non-thematic conditions; recall of consistent features was greater in the non-thematic/diagnostic condition than in the non-thematic/non-diagnostic condition, $t(28) = 2.44$, $p = .021$.

A question of particular interest concerned the relative sizes of the effects of adding diagnostic labels versus thematic relevance:

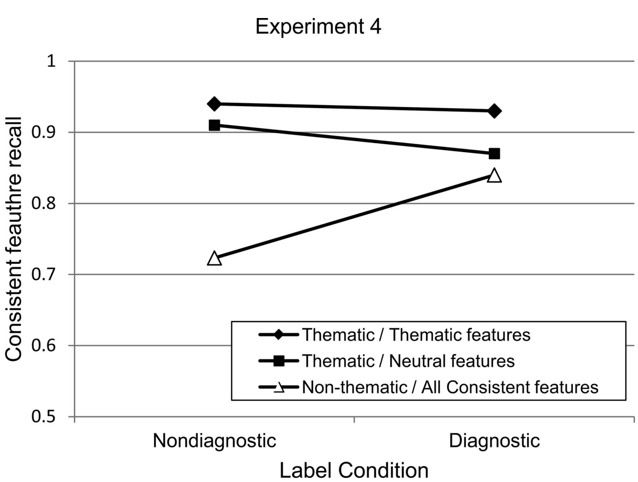

Figure 5. Consistent feature recall data from Experiment 4, averaged over the second half of trials. The top line shows data for thematic features of the two thematic conditions (non-diagnostic vs. diagnostic), the middle line shows data for neutral features from the same two conditions, and the bottom line shows data for all consistent features within the two non-thematic conditions (nondiagnostic vs. diagnostic).

Table 7
Recall Memory Data From Experiment 4

Condition	Dependent variable	*M*
Thematic/diagnostic	Consistent feature recall (thematic)	.93
	Consistent feature recall (neutral)	.87
	Variable feature recall	.68
Thematic/non-diagnostic	Consistent feature recall (thematic)	.94
	Consistent feature recall (neutral)	.91
	Variable feature recall	.59
Non-thematic/diagnostic	Consistent feature recall (pooled)	.84
	Variable feature recall	.65
Non-thematic/non-diagnostic	Consistent feature recall (pooled)	.72
	Variable feature recall	.53

Which manipulation produced the greater improvement compared to a non-thematic/non-diagnostic condition? Although diagnostic labels improved recall in the non-thematic conditions, recall of consistent dimensions in the non-thematic/diagnostic condition remained significantly below that of thematic dimensions in the pooled thematic conditions, $t(43) = 2.79$, $p = .008$; there was no corresponding difference in recall of neutral dimensions, $t(43) = 1.24$, $p = .221$. Thus, thematic relevance had a greater effect than diagnostic labels on recall of thematic (but not neutral) dimensions. Moreover, adding diagnostic labels to an already thematic condition produced no additional increase in performance. Within the two thematic conditions, thematic dimensions were recalled significantly better than neutral dimensions, $t(29) = 2.60$, $p = .015$.

The variable dimensions showed no effect of theme and no significant interaction, $F(1, 56) = 1.10$, $p = .298$ and $F(1, 56) = 0.123$, $p = .727$, respectively. The effect of label was significant, however, $F(1, 56) = 6.01$, $p = .017$, with asymptotic recall in the diagnostic conditions averaging .67 and that in the non-diagnostic condition averaging .56. Recall of the instance labels was not affected significantly by any factor, with asymptotic recall averaging about .82.

Discussion

The first major result of this experiment was that adding diagnostic verbal labels to a non-thematic baseline condition improved category learning, as reflected in a 12% increase in consistent feature recall. The second result was that adding thematic relevance also improved learning. This improvement was greater than the diagnostic label effect (about 22% for thematic dimensions, almost double the size of the labeling effect). The third major result was the interaction (non-additivity) between these two factors (adding diagnostic labels to a thematic condition produced no further improvement in learning).

These results are consistent with the CBS model presented earlier, in which the effect of thematic relevance reflects both category cuing and feature facilitation (i.e., total knowledge effect = category cuing effect + feature facilitation effect). This model assumes that the themes provide a cue that helps people to recognize separate categories and in addition facilitates the learning of thematically relevant features of those categories. Because the themes already provide an efficient category cue, the model predicts that adding diagnostic labels would provide no additional benefit, consistent with the present results. The model also predicts that the thematic effect, consisting as it does of both category cuing and feature facilitation, should be larger than the diagnostic label effect, which consists of category cuing only. This prediction only applies to thematic features; as thematic relevance would not directly facilitate learning of neutral features, the thematic effect should be the same as the labeling effect for these features. Again, these predictions are confirmed by the data: The thematic conditions showed higher recall than the non-thematic/diagnostic condition for thematic dimensions, but there was no difference for neutral dimensions.

Experiment 5

The experiments reported so far have provided strong evidence for the distinction between category cuing and feature facilitation effects proposed at the beginning of this article. However, they have provided little evidence concerning the nature of the feature facilitation effect itself, that is, the extent to which this effect is a direct facilitation (binding) effect, an indirect, release-from-interference (segregation) effect, or some combination of the two. As noted in the introduction, Clapper (2007) reported data suggesting that the feature facilitation effect in the exemplar-memory task is largely, if not entirely, a segregation effect rather than a binding effect. The experiment described next is another attempt to disentangle and measure these two potential sources of feature facilitation.

Two of the previous experiments focused on comparing same-theme to opposite-theme conditions. So far, the evidence seems to suggest that knowledge effects in the same-theme conditions are weak or nonexistent. The reason for this lack of effect, it has been argued, is that relating two new categories to the same familiar theme provides no category cuing benefit. It follows that, if people are not inventing separate categories, thematic knowledge would have no real way to facilitate the learning of individual features within those categories. Hence, neither category cuing nor feature facilitation effects would be expected in a same-theme condition.

The outcome might be quite different if people were provided with an effective category cue in a same-theme condition (e.g., if diagnostic verbal labels were provided, as in Experiment 4). In this case, people would presumably succeed in discovering separate categories (due to the labels), and having relevant prior knowledge might then facilitate learning the features within those categories. Importantly, any such feature facilitation would of necessity be a feature binding effect, for the simple reason that a common theme would provide no discriminating information for a feature segregation effect.

If the assumption that any feature facilitation effect in the same-theme/diagnostic label condition would have to be a binding effect is granted, the size of the knowledge effect in this condition relative to that in an opposite-theme condition can be used to assess the relative importance of binding versus segregation in the present task. The first comparison needed would be between the same-theme condition and a no-theme baseline condition (with diagnostic labels); to the extent that there is a significant feature binding effect in the same-theme condition, memory for thematic features should be higher than that of the corresponding features in the no-theme/diagnostic condition and higher than that of neutral consistent features within the same-theme condition itself. The second comparison would be between same-theme and opposite-theme conditions; to the extent that feature facilitation is stronger in the latter than in the former, the difference would necessarily be due to the extra segregation effect in the opposite-theme condition (and the size of the difference would allow a rough estimate of the size of this effect).

In summary, the main objectives of this experiment are to (a) provide an effective category cue in each condition, thereby "subtracting out" differences in category cuing and allowing one to (b) estimate the relative importance of feature binding versus segregation as components of the feature facilitation effect, by comparing the size of this effect in same-theme versus opposite-theme conditions. The experiment included two conditions similar to the thematic/diagnostic and non-thematic/diagnostic conditions from Experiment 4, here referred to as opposite-theme versus no-theme conditions. In addition, a third condition similar to the same-theme condition of Experiment 3 was included, the only difference being that diagnostic verbal labels were included in this condition, as in the opposite-theme and no-theme conditions.

Method

Participants. Sixty-four undergraduate students of California State University, San Bernardino, participated in exchange for extra credit in a psychology class of their choice.

Materials, procedure, and design. The procedure and materials were identical to those of Experiment 4 and earlier studies. Participants were randomly assigned to three conditions, referred to as the opposite-theme, same-theme, and no-theme conditions. The opposite-theme condition was identical to the thematic/diagnostic condition from Experiment 4, and the no-theme condition was identical to the non-thematic/diagnostic condition from the same experiment. The same-theme condition was similar to the condition of the same name in Experiments 1 and 3; it was also identical to the present opposite-theme condition except that the thematic values of the two categories were from the same rather than opposite themes (e.g., two categories of senior citizens rather than senior citizens vs. adolescents). The main difference between this experiment and Experiment 3 (which also compared opposite-theme, same-theme, and no-theme conditions) is the presence of diagnostic labels in all conditions of the present experiment.

Results

The main data relevant to evaluating the experimental hypotheses were recall of thematic and neutral consistent features compared across the three conditions. The data are shown over trials in Figure 6; condition means are shown in Table 8 and statistical tests are shown in Table 9.

A one-way ANOVA showed an overall effect of conditions on thematic feature memory, $F(2, 26) = 4.57$, $p = .015$, but no effect on memory for neutral features, $F(2, 26) = 1.17$, $p = .322$; variable features, $F(2, 26) = 0.058$, $p = .579$; or verbal labels, $F(2, 46) = 1.89$, $p = .164$. Planned comparisons showed that memory for thematic dimensions was significantly greater in the opposite-theme condition than in the same-theme and no-theme conditions but that there was no significant difference between the latter two conditions (see Table 9). Consistent with the nonsignificant ANOVA, all t tests between conditions were nonsignificant for neutral dimensions.

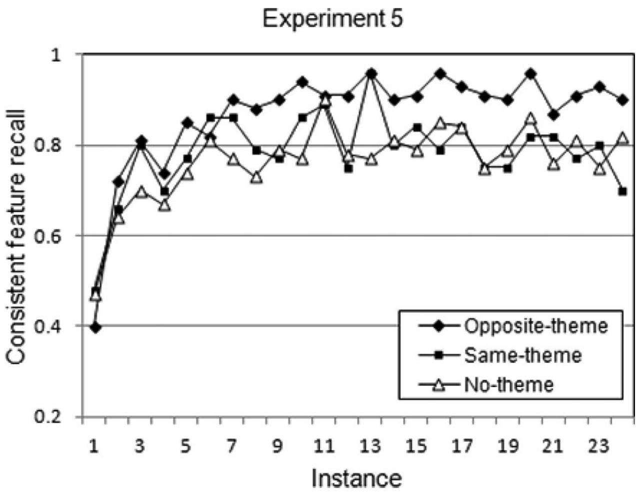

Figure 6. Consistent feature recall data from Experiment 5, plotted over trials. The upper line shows data from thematic features within the opposite-theme condition, the lower line with dark squares shows data from thematic features within the same-theme condition, and the lower line with white triangles shows data averaged over all consistent features within the no-theme condition.

There were no significant differences between the no-theme and same-theme conditions in this experiment, so they were pooled in further analyses comparing them to the opposite-theme condition. Although the difference between opposite-theme and the two other conditions remained significant for thematic dimensions, it remained nonsignificant for neutral dimensions (see Table 9). In addition, the fact that the opposite-theme and no-theme conditions in the present experiment were identical to the thematic/diagnostic and non-thematic/diagnostic conditions from Experiment 4 allows some comparisons to be made across the two experiments. Pooling the data from the two experiments again yields a significant difference between the opposite-theme and no-theme conditions for thematic dimensions but not for neutral dimensions (see Table 9). Thus, the lack of effect for neutral dimensions cannot be attributed to a simple lack of statistical power.

Discussion

This experiment attempted to eliminate differences in category cuing by including diagnostic labels in all three conditions, thereby allowing feature facilitation to be examined in relative isolation. As expected based on the results of Experiment 4, a robust knowledge effect was found in the opposite-theme condition compared to the no-theme condition. Assuming that people recognize separate categories in both conditions due to the diagnostic labels, thus eliminating differences in category cuing, this difference must be a feature facilitation effect. This effect occurred only in the opposite-theme condition and not in the same-theme condition. This pattern of results was consistent with a feature segregation effect dependent on the presence of contrasting themes and provided no evidence for any binding effects within individual categories.

Importantly, the benefits observed in the opposite-theme condition relative to the no-theme condition were observed for thematic features only; there was no significant difference between these two conditions in memory for neutral consistent features. This replicates the lack of difference between the same conditions in Experiment 4 and is consistent with the hypothesis that the neutral-feature benefit is due to category cuing and that when differences in category cuing are removed (by adding diagnostic labels to both conditions), differences in neutral feature recall disappear as well.

General Discussion

The results of these five experiments are readily interpretable in terms of the CBS model. In this section, I summarize their major results and describe their implications from the model's perspective. I then discuss the implications of these results for research in several areas related to unsupervised category learning.

Summary of Major Results

1. A robust knowledge effect was consistently obtained for both thematic and neutral features when performance in opposite-theme conditions was compared to that in no-theme baseline conditions. Within the CBS framework, such facilitation for all consistent features is expected due to category cuing effects. This is based on the assumption that noticing the themes would cause people to recognize separate categories, which in turn would make it possible for them to learn the consistent features associated with those categories. The result would be better recall of all consistent features (both thematic and neutral), as found in the present experiments.

2. The benefits of prior knowledge were greater for thematically relevant features than for neutral features (i.e., thematic features were generally better recalled than neutral features in opposite-theme conditions). This is consistent with a feature facilitation effect, based on the assumption that knowledge should be more helpful in learning features that are related to that knowledge than features that are not.

3. Diagnostic verbal labels were effective category cues, as evidenced by the fact that adding diagnostic labels to a non-thematic condition produced a significant improvement in learning. (Diagnostic labels had been shown previously to improve learning in a similar task using forced-choice recognition rather than cued recall, Clapper & Bower, 2002.)

4. There were no consistent knowledge benefits in the same-theme conditions, compared to equivalent no-theme conditions. The lack of benefits in the combined data from Experiments 1 and 3, in which diagnostic labels were not provided, is consistent with

Table 8
Recall Memory Data From Experiment 5

Condition	Dependent variable	*M*
Opposite-theme	Consistent feature recall (thematic)	.92
	Consistent feature recall (neutral)	.87
	Variable feature recall	.65
Same-theme	Consistent feature recall (thematic)	.80
	Consistent feature recall (neutral)	.82
	Variable feature recall	.63
No-theme	Consistent feature recall (pooled)	.80
	Variable feature recall	.59

Table 9
Between-Groups Planned Comparisons From Experiment 5

Type of dimension/condition	*t*	*df*	*SE*	*p*
Thematic dimensions				
Opposite- vs. same-theme	2.44	29	.047	.021
Opposite- vs. no-theme	3.11	33	.038	.004
Same- vs. no-theme	0.52	30	.050	.959
Opposite- vs. others (pooled)	3.06	47	.038	.004
Opposite- vs. no-theme (combined with Experiment 4)	3.61	63	.029	.001
Neutral dimensions				
Opposite- vs. same-theme	0.89	29	.057	.382
Opposite- vs. no-theme	1.62	33	.046	.116
Same- vs. no-theme	0.45	30	.052	.654
Opposite- vs. others (pooled)	1.47	47	.043	.149
Opposite- vs. no-theme (combined with Experiment 4)	1.55	63	.033	.127
Variable dimensions				
Opposite- vs. same-theme	0.33	29	.063	.742
Opposite- vs. no-theme	1.04	33	.062	.306
Same- vs. no-theme	0.67	30	.064	.508
Opposite- vs. others (pooled)	0.85	47	.045	.402

the model's assumption that benefits in the opposite-theme conditions depended on knowledge increasing the probability that people would create separate categories (i.e., that it would have a category cuing effect). However, knowledge effects were also not observed in the same-theme condition of Experiment 5, in which people were given diagnostic labels to serve as category cues. Thus, even when separate categories were given by the experimenter, knowledge provided no additional feature facilitation if both categories were related to the same theme. This is consistent with previous results (Clapper, 2007), suggesting that the feature facilitation observed in this task is primarily a segregation effect rather than a binding effect.

5. Even when diagnostic labels were provided to both (in Experiment 5), learning in the opposite-theme condition was greater than that in the no-theme condition but only for thematic features; there was no corresponding difference for neutral features. In this comparison, providing labels in both conditions was intended to subtract out differences in category cuing and leave only differences in feature facilitation. Thematic features were expected to benefit from this facilitation but neutral features were not, consistent with the actual results.

6. The effect of thematic relevance was significantly greater than that of diagnostic labels, and the two effects were nonadditive. That is, thematic relevance and diagnostic labels together (the thematic/diagnostic condition from Experiment 4) were no better than thematic relevance alone (the thematic/non-diagnostic condition, see Figure 5). This is consistent with the model's assumption that the effect of adding thematic knowledge is the sum of the diagnostic label effect (i.e., category cuing) plus the feature facilitation effect. Given this assumption, adding diagnostic labels to a thematic condition would be redundant and would provide no new information to the learner (analogous to providing the learner with the same label twice).

Remaining Questions About the Results

Why was there no evidence for feature binding effects? In the introduction, it was noted that previous experiments (Clapper, 2007) had also failed to provide evidence for binding effects, and it was suggested that such effects are most likely to occur when the person's ability to generate the correct response is the key factor limiting memory performance. On the other hand, if all responses (including the correct response) are already at ceiling availability, there would be no room for further activation of the correct response to have a net effect in improving performance. In a forced-choice recognition task such as that used in Clapper (2007), all possible response alternatives are presented explicitly at the time of test. Therefore, response generation is not a limiting factor, and the onus is on the person's ability to distinguish correct from incorrect responses (e.g., by inhibiting or eliminating the incorrect responses). This explains why segregation effects loomed so much larger than binding effects in those experiments.

In the present experiments, a cued recall task was employed in place of the recognition task used in Clapper (2007). One potential advantage of this switch is that recall tasks place a greater emphasis on generating the correct response from memory, as it is not presented as one of the choice alternatives at test. Moreover, the incorrect alternatives are also not presented for each test, which might tend to decrease their availability and make the discrimination aspect of the memory task less challenging. For these reasons, the recall task might be expected to provide greater opportunities than the recognition test for detecting a significant binding effect.

The results of the present experiments, however, provided no more evidence for binding effects than did those of Clapper (2007). One interpretation of this outcome is that the change from recognition to recall did not alter the overall dominance of discrimination over generation in the present task. Indeed, such dominance might hold for most situations in which people are learning artificial categories constructed from a limited set of discretely varying attributes. In such cases, all possible values of a given attribute are likely to become readily available from memory as the person becomes familiar with the task and stimulus set, whether or not these values are presented as explicit alternatives on the memory tests. If so, the key factor determining performance would be the person's ability to discriminate which value was associated with the current category, meaning that the main way in which prior knowledge could facilitate recall would be via segregation effects.

This suggests that segregation effects are likely to dominate whenever people are learning artificial, discretely varying stimulus sets. The same would not necessarily be true in naturalistic settings, where the possible values of a given attribute are rarely limited to a small, fixed number of highly available alternatives. In such cases, in which potential responses start out with low baseline availability, increasing the availability of the correct response might well increase the likelihood that the person would generate that response and hence increase recall performance. Thus, despite the lack of positive evidence so far, the present results cannot be taken as blanket proof that feature binding effects do not exist or are unimportant in real-world settings.

Why do variable as well as consistent attributes show knowledge effects? In most of the experiments reported above, variable as well as consistent dimensions were recalled better in thematic than in non-thematic conditions. This result appears counterintuitive at first blush, because these features vary unpredictably from instance to instance and so cannot benefit directly from category knowledge. However, improved memory for variable as well as consistent features is actually a predicted outcome of discovering separate categories in the exemplar-memory task (Clapper, 2007; Clapper & Bower, 2002). According to the category invention model, after discovering separate categories people should find the consistent features of successive instances increasingly familiar and predictable and hence progressively easier to recall. As this occurs, there is less need to spend time attending to these features during the study phase of each trial and hence more time available to study the variable attributes. (Another way to say this is that the variable attributes, being less predictable than the consistent attributes, provide more information per unit of study time invested.) The expected result would be improved memory for both variable and consistent attributes in this task.

Previous studies (Clapper, 2006; Clapper & Bower, 2002) have directly tested this prediction about declining study times to consistent features and increasing study times to variable features. This was done by presenting the features of each training instance in a display in which all the features but one were blocked out by rows of Xs. As they scanned up and down the list, the person could uncover and study only one feature at a time, allowing the computer to record how much time they spent studying each feature. As expected, people who learned the categories tended to allocate less time to consistent than to variable features and to show improved memory for both.

Was the improved recall of consistent features due to guessing or memory? In principle, people have two options when they attempt to recall the features of a specific instance: They can consult their specific memory trace for that instance, or failing that, they can attempt to use their general knowledge about its category to guess or infer the correct features. In the case of consistent attributes, they can achieve perfect performance by simply filling in the expected values, because these are the only values ever presented within a given category in the present experiments. Thus, the improved recall of consistent features was presumably due to category-based inference (guessing), not to a direct improvement in memory for the features of specific instances. This is consistent with the study time results discussed above, which showed that people actually spent less time studying consistent than variable attributes and so probably recorded weaker associations between these features and the specific instances in which they occurred.

On implication of this strategy is that people should show poor memory discrimination for the consistent features of individual instances. Indeed, this is a standard result in experiments on memory for knowledge-based material. For example, Graesser and colleagues have carried out numerous studies of people's memory for short textual descriptions (prose passages) based on common event categories, or scripts (see, e.g., Graesser, 1981; Graesser, Woll, Kowalski, & Smith, 1980; Smith & Graesser, 1981). They found that people had much poorer memory discrimination for typical script-based events than for atypical or script-irrelevant events. In particular, people were highly accurate in recognizing or recalling script-typical events when they were included in the passage but were just as likely to falsely remember them when they were not. Similar results have been obtained in other domains. For example, Stangor and McMillan (1992) reported a meta-analysis of over two dozen studies of social stereotype effects on recognition memory. They found a very consistent pattern in which atypical or unexpected items showed higher memory discriminability than typical or expected items, largely due to people's tendency to falsely recognize the latter type of item.

This comparison suggests that people should also show low discrimination (d') for consistent features in the present task, as they should be equally likely to recall such features whether or not the features actually occurred in a particular instance. However, the present stimuli differed from the textual stimuli used in the research described above in consisting of substitutive rather than additive features (Gati & Tversky, 1984), with these features listed in same order on every trial. In this case, omitting an expected value implies replacing it with an unexpected value, which should be regarded as surprising and "tagged" as an expectation failure in memory (e.g., Schank, 1982). This strategy of ignoring consistent values when present but explicitly noting when they are absent or replaced by exceptional values could result in high measured discriminability (d') based on guessing rather than on actual memory. Indeed, the general finding of low d' for expected features might be restricted to textual stimuli (prose passages), which readers expect to be elliptical and to omit important information. The same might not be true of real objects, dimensionalized lists, and other stimuli where the absence of expected features would be informative and easily noticed.

Implications for Related Research

Knowledge effects on category construction. The CBS model was intended as a framework for predicting knowledge effects in the exemplar-memory task. Another task used to study unsupervised learning is category construction, in which participants sort a set of stimuli into two categories in any way that appears natural to them. Prior knowledge facilitates performance in this task; specifically, people are more likely to sort stimuli into categories based on correlational structure (family resemblance) if the features of the two categories are related to familiar traits or themes or if some ad hoc coherence can be discovered to relate them (e.g., Kaplan & Murphy, 1999; Medin et al., 1987; Spalding & Murphy, 1996). In the absence of such organizing themes, people tend to sort the stimuli into categories based on a single dimension and show no awareness of the stimulus set's correlational structure. This lack of sensitivity is reminiscent of that found in the present experiments, in which people without themes or diagnostic labels also showed little learning of correlational structure.

A series of experiments by Kaplan and Murphy (1999) provides an illustration of how prior knowledge can affect performance in this task. In these experiments, people were more likely to create categories consistent with the correlational structure of a set if each instance had a single idiosyncratic feature that could be related to a common theme. For example, different vehicles might be described as driving on glaciers, having treads, and being heavily insulated, all of which can be related to a common "arctic vehicle" theme. (Other stimuli in the same set would have features relating to a contrasting "jungle vehicle" theme.) Importantly, the presence of this single thematic feature improved people's learning of other, non-thematic features within a category. Data from later inference tests, in which people attempted to judge which pairs of features occurred together during the previous training phase, confirmed that correlations among neutral features were learned better when a thematic feature was present and diagnostic of the correlational structure of the set. Kaplan and Murphy suggested that the thematic-idiosyncratic features prompted people to relate the non-thematic features with each other and with the thematic features; this elaboration process resulted in better learning of all features.

Another explanation considered by Kaplan and Murphy (1999) assumes that people perceived the idiosyncratic features as expressions of a single thematic dimension ("vehicle type") with two alternative values (arctic, jungle). This higher order dimension could function as an implicit category label, providing the basis for a one-dimensional sort of the instances into two theme-based categories. By partitioning the set based on the themes and relating the correlated features to these themes, the number of correlations to be learned would be greatly reduced (see introduction), resulting in better learning of both thematic and neutral features. Kaplan and Murphy attempted to test this explanation by showing examples of two family resemblance categories in different colors during an initial familiarization phase (prior to testing). In principle, the color coding manipulation should provide a highly salient cue to category membership, analogous to that provided by an implicit theme. However, this manipulation did not improve performance during a later test phase, leading them to reject the implicit label/category cuing hypothesis as an explanation for their data.

Of course, the mere fact that a highly salient feature (like color) is diagnostic of category membership is no guarantee that people will notice this fact or use it to divide the stimuli into categories. Indeed, nine features of every instance were perfectly correlated with (diagnostic of) category membership in the present experiments, but in the absence of labels or themes people rarely seemed to notice this. It is possible that if the instances in Kaplan and Murphy's (1999) experiment had been presented with diagnostic category labels rather than being shown in different colors during the familiarization phase, these labels might had the same effect as the thematic features (i.e., they may have caused people to later sort the instances into categories consistent with the correlational structure of the set). In this case, improved learning of correlated features would follow as a consequence of correctly partitioning the instances, as explained above.

Another way to distinguish the category cuing versus within-category elaboration hypotheses would be to compare a same-theme condition to a different-theme condition, as in some of the present experiments. If the idiosyncratic features were all related to the same theme across both categories (i.e., all related to jungle vehicles rather than half to jungle vehicles and half to arctic vehicles), they could in principle still support the kind of within-category elaboration process envisioned by Kaplan and Murphy (1999). However, in this case the idiosyncratic features would not provide a useful cue to help people distinguish separate categories, so the CBS model would predict no benefit. If a significant facilitation of sorting and/or inference testing performance was observed in the opposite theme but not in the same-theme condition, this would support a category cuing explanation for Kaplan and Murphy's earlier results.

Knowledge effects on inference learning. Another type of task to which the current results may be relevant is *inference learning* (e.g., A. L. Anderson, Ross, & Chin-Parker, 2002; Chin-Parker & Ross, 2004; Yamauchi & Markman, 1998). In a typical inference learning experiment, the participant is shown a series of labeled training instances, each missing a single feature. Two or more alternative versions or values of the missing feature are shown; the participant's task is to select the correct value for the current instance. In some experiments, instances with the same category label are divisible into subcategories based on within-category correlational structure, similar to the way in which categories are defined in the present experiments. Any learning of these subcategories would necessarily be unsupervised, as no labels or feedback are provided to indicate their existence to the participant.

Most of the research on inference learning has focused on comparing this task to supervised classification (e.g., Chin-Parker & Ross, 2004; Jones & Ross, 2011), and there has been little investigation of the effects of prior knowledge on inference learning itself. However, Rehder and Ross (2001) and Erickson, Chin-Parker, and Ross (2005) each reported several studies investigating how inference learning is affected by the abstract coherence of the categories being learned. This coherence was more like an ad hoc theme or scenario that the person had to recognize on the spot (Barsalou, 1983) than the type of already familiar trait themes used in the present experiments. However, the results were similar, in that people performed much better on the inference learning task when the unlabeled subcategories (defined by correlational structure) were aligned with (predictable from) the abstract themes or coherence.

One way to explain these results in terms of the CBS framework would be to assume that the abstract coherence, once noticed, helps people sort the stimuli into appropriate subcategories, and this in turn helps them fill in the correct values on the inference tests—in other words, that the coherence has a category cuing effect. Consistent with this, participants' verbal responses to debriefing questions posed by Rehder and Ross (2001) showed that participants recognized distinct subtypes of the target category in the coherent condition (consistent with their actual correlational structure) but not in the incoherent condition (in which the same degree of correlational structure was present but was not predictable based on abstract coherence). As described earlier, the ad hoc themes might have been perceived as a kind of emergent or higher order feature that people could use to partition the sets into categories, making it possible to learn their correlational structure.

There are admittedly important differences between these different procedures, complicating any comparisons between them. (For example, the coherence used in the inference learning experiments was more abstract and ad hoc than the familiar trait themes used in the present experiments.) Nevertheless, it does seem possible to design inference learning experiments analogous to the exemplar-memory studies described here. The main procedural difference is that people in inference learning tasks attempt to predict (rather than recall) individual features of specific instances, with the rest of the features present as cues at test. As in the present experiments, the stimuli could be designed to fall into separate categories based on correlational structure, and these could be related to familiar themes and/or provided with diagnostic labels. One reason to carry out such comparative investigations would be to evaluate how prior knowledge affects learning in different types of tasks. For example, people in inference learning tasks typically place greater emphasis on learning the internal structure of categories than those engaged in supervised classification (e.g., Chin-Parker & Ross, 2004; Jones & Ross, 2011). This suggests that people might show stronger feature binding effects in inference learning than they have so far in the exemplar-memory task (e.g., stronger knowledge effects in a same-theme condition when diagnostic labels are provided).

Knowledge effects on supervised classification learning. The task most often used to study category learning is supervised classification, in which people assign training instances to predefined categories based on experimenter feedback. Because participants are told which category each instance belongs to, there is no real opportunity for prior knowledge to play a category cuing role in this task; hence, the main effect of prior knowledge would have to be some form of feature facilitation. As for the nature of this feature facilitation effect, most theorists have focused on how knowledge can increase within-category coherence (e.g., through causal elaboration), leading to better learning of feature-to-feature and/or feature-to-category connections (e.g., Kaplan & Murphy, 1999; Murphy & Medin, 1985). Without denying the basic plausibility of this idea, it nonetheless makes sense to inquire to what extent such within-category connections improve performance directly, by priming or activating correct responses, versus indirectly, by making it easier to reject or eliminate incorrect responses.

In the context of the present discussion, the most obvious way to frame this issue is in terms of the same-versus-opposite-theme comparisons used in the present experiments; in particular, it might be interesting to employ the same manipulation in a supervised task. For example, a set of experiments by Wattenmaker et al. (1986) showed that people learned categories based on probabilistic correlational structure more quickly if their features were consistent with contrasting themes. However, it is possible to imagine a learning process—especially one in which people engage in active, explanation-driven processing to link the features within a concept—that would expect essentially the same result whether the categories were related to different or to the same themes. But even if no knowledge effects at all were found in a same-theme version of this task—as none were found in the present experiments—this would not imply that knowledge does not bind or link the features within a category (indeed, it is difficult to imagine how it could do otherwise). Rather, the results would allow one to determine whether the effects of this linkage on actual performance are due to one factor (enhanced discrimination) or to another (say, greater activation of the target concept at the time of test).

One issue that complicates attempts to study knowledge effects in supervised learning is the fact that improved ability to identify instances of one category generally implies improved ability to identify the other category, as well. Due to the forced-choice nature of this task, if only one of two categories in a set is related to prior knowledge, both are likely to benefit by default. This problem is less obvious in single-feature classification tests, in which features are presented individually and participants are asked to judge which category a given feature most likely implies. (With this testing format, it is easy to imagine that people might have higher average accuracy for the features of one category than those of the other.) Single-feature tests also make it possible to compare learning of thematically relevant versus irrelevant features within the same category. In a study by Heit and Bott (2000), for example, people learned two knowledge-related categories containing both thematically relevant and thematically neutral features. Single-feature classification performance was greater for relevant than for neutral features, the basic feature facilitation effect expected within the CBS framework. Further studies using this procedure combined with some of the current manipulations (e.g., same vs. opposite themes) might provide a way to investigate the nature of this effect.

Relevance for Computational Modeling

As already noted, existing computational models of knowledge effects on category learning are generally specific to supervised learning (e.g., Heit & Bott, 2000; Rehder & Murphy, 2004; the integration model of Heit, 1994, 1998, is a partial exception), and none are directly applicable to the present exemplar-memory task. Although these models cannot be used to provide accounts of the present data, they are relevant to the present discussion in at least two ways. First, some of the manipulations employed in the present experiments and those of Clapper (2007), particularly the comparison between same- versus opposite-theme conditions and the effects of having only one category in a set related to prior knowledge, could be employed in a supervised learning task. This means that simulation studies could be set up using the models in question and the results compared to the data from actual participants. Such studies might provide useful information for extending and improving the models.

A second way in which these models are relevant to the present discussion is that they may provide the basis for future computational theories that can accommodate a wider variety of category learning tasks, including the present exemplar-memory task. One major feature that should be added to the models would be some mechanism for discovering separate categories when these are not provided by the experimenter. Additional mechanisms would be needed to model memory for individual features in association with specific instances while continuing to account for standard performances such as whole-instance and single-feature classification. Fully accounting for the present results would almost certainly require substantial modification and elaboration of existing models.

Summary and Conclusions

Taken together, the experiments described in this article provide strong evidence that prior knowledge can facilitate unsupervised learning in at least two ways: first, by helping people to discover separate categories within a given training set and, second, by facilitating their learning of individual features within those categories. The failure to obtain clear binding effects suggests that much of the feature facilitation effect—at least in the present task—may actually be due to enhanced segregation and reduced interference between the features of different categories rather than, say, better cuing due to stronger associations among features within a category. The implications of these results were discussed for other types of category learning tasks and for existing computational models of knowledge effects on learning. It appears that the CBS model may have interesting applications to other tasks involving unsupervised learning, as well as suggesting potentially interesting manipulations and theoretical distinctions pertaining to supervised learning. On a more general level, the present results demonstrate how knowledge can play multiple roles in human learning, depending on the nature of the task and the type of information being acquired. Research that attempts to carefully distinguish different types of knowledge effects across a variety tasks and situations hopefully will continue to illuminate these roles in the years ahead.

References

Anderson, A. L., Ross, B. H., & Chin-Parker, S. (2002). A further investigation of category learning by inference. *Memory & Cognition, 30,* 119–128. doi:10.3758/BF03195271

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory.* Washington, DC: Winston.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11,* 211–227. doi:10.3758/BF03196968

Chin-Parker, S., & Ross, B. H. (2004). Diagnosticity and prototypicality in category learning: A comparison of inference learning and classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 216–226. doi:10.1037/0278-7393.30.1.216

Clapper, J. P. (2006). When more is less: Negative exposure effects in unsupervised learning. *Memory & Cognition, 34,* 890–902. doi:10.3758/BF03193435

Clapper, J. P. (2007). Prior knowledge and correlational structure in unsupervised learning. *Canadian Journal of Experimental Psychology, 61,* 109–127. doi:10.1037/cjep20070012

Clapper, J. P., & Bower, G. H. (1991). Learning and applying category knowledge in unsupervised domains. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 27, pp. 65–108). New York, NY: Academic Press.

Clapper, J. P., & Bower, G. H. (1994). Category invention in unsupervised learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 443–460. doi:10.1037/0278-7393.20.2.443

Clapper, J. P., & Bower, G. H. (2002). Adaptive categorization in unsupervised learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 908–923. doi:10.1037/0278-7393.28.5.908

Erickson, J. E., Chin-Parker, S., & Ross, B. H. (2005). Inference and classification learning of abstract coherent categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 86–99. doi:10.1037/0278-7393.31.1.86

Gati, I., & Tversky, A. (1984). Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology, 16,* 341–370. doi:10.1016/0010-0285(84)90013-6

Graesser, A. C. (1981). *Prose comprehension beyond the word.* New York, NY: Springer-Verlag.

Graesser, A. C., Woll, S. B., Kowalski, D. J., & Smith, D. A. (1980). Memory for typical and atypical actions in scripted activities. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 503–515. doi:10.1037/0278-7393.6.5.503

Heit, E. (1994). Models of the effects of prior knowledge on category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1264–1282. doi:10.1037/0278-7393.20.6.1264

Heit, E. (1998). Influences of prior knowledge on selective weighting of category members. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 712–731. doi:10.1037/0278-7393.24.3.712

Heit, E., & Bott, L. (2000). Knowledge selection in category learning. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 39, pp. 163–199). San Diego, CA: Academic Press.

Jones, E. L., & Ross, B. H. (2011). Classification versus inference learning contrasted with real-world categories. *Memory & Cognition, 39,* 764–777. doi:10.3758/s13421-010-0058-8

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93,* 136–153. doi:10.1037/0033-295X.93.2.136

Kaplan, A. S., & Murphy, G. L. (1999). The acquisition of category structure in unsupervised learning. *Memory & Cognition, 27,* 699–712. doi:10.3758/BF03211563

Keppel, G. (1968). Retroactive and proactive inhibition. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 172–213). Englewood Cliffs, NJ: Prentice Hall.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19,* 242–279. doi:10.1016/0010-0285(87)90012-0

Michalski, R. S., & Stepp, R. E. (1983). Learning from observation: Conceptual clustering. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (pp. 331–364). Palo Alto, CA: Tioga.

Murphy, G. L. (2002). *The big book of concepts.* Cambridge, MA: MIT Press.

Murphy, G. L., & Allopenna, P. D. (1994). The locus on knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 904–919. doi:10.1037/0278-7393.20.4.904

Murphy, G. L., & Kaplan, A. S. (2000). Feature distribution and background knowledge in category learning. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53*(A), 962–982.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289–316. doi:10.1037/0033-295X.92.3.289

Pazzani, M. J. (1991). Influence of prior knowledge on concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 416–432. doi:10.1037/0278-7393.17.3.416

Postman, L. (1971). Transfer, interference, and forgetting. In J. W. Kling & L. A. Riggs (Eds.), *Experimental psychology* (3rd ed., pp. 1019–1132). New York, NY: Holt, Rinehart & Winston.

Rehder, B., & Murphy, G. L. (2003). A knowledge-resonance (KRES) model of category learning. *Psychonomic Bulletin & Review, 10,* 759–784. doi:10.3758/BF03196543

Rehder, B., & Ross, B. H. (2001). Abstract coherent categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1261–1275. doi:10.1037/0278-7393.27.5.1261

Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814. doi:10.1037/0278-7393.21.4.803

Schank, R. C. (1982). *Dynamic memory.* Cambridge, England: Cambridge University Press.

Smith, D. A., & Graesser, A. C. (1981). Memory for actions in scripted activities as a function of typicality, retention interval, and retrieval task. *Memory & Cognition, 9,* 550–559. doi:10.3758/BF03202349

Spalding, T. L., & Murphy, G. L. (1996). Effects of background knowledge on category construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 525–538. doi:10.1037/0278-7393.22.2.525

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin, 111,* 42–61. doi:10.1037/0033-2909.111.1.42

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology, 18,* 158–194. doi:10.1016/0010-0285(86)90011-3

Wisniewski, E. J. (1995). Prior knowledge and functionally relevant features in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 449–468. doi:10.1037/0278-7393.21.2.449

Yamauchi, T. (2005). Labeling bias and categorical induction: Generative aspects of category information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 538–553. doi:10.1037/0278-7393.31.3.538

Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory and Language, 39,* 124–148. doi:10.1006/jmla.1998.2566

Appendix

Stimulus Materials for the Experiments

Dimension	Values
	Highbrow/lowbrow theme
1. favorite drink	sherry; fine wine; beer; coffee; cola; milk; juice
2. favorite activity	yachting; polo; bowling; soccer; softball; cycling; fishing
3. enjoys watching	the opera; the symphony; pro wrestling; basketball; movies; baseball; television
4. drives	a Mercedes; a Lamborghini; an old pickup; a Honda; a Toyota; a Subaru; a Volvo
5. lives in	Beverly Hills; Malibu; Detroit; Sacramento; Tucson; Eugene; Santa Fe
6. employed as	a lawyer; a plastic surgeon; is unemployed; an accountant; a technician; a manager; a consultant
7. favorite food	Russian caviar; escargot; pizza; steak; fish; pork; oysters
8. last vacation	Paris; Tuscany; Tijuana; New Orleans; Disney World; San Francisco; New York City
9. favorite music by	Mozart; Bach; Aerosmith; the Beatles; BB King; Aretha Franklin; Janis Joplin
10. clothes by	Armani; Versace; Discount Mart; Old Navy; The Gap; Levi's; Gottschalks
11. favorite TV show	*Masterpiece Theater*; *NOVA*; *The Jerry Springer Show*; the evening news; *ER*; *Monk*; *Friends*
12. graduated from	Harvard; Princeton; high school; community college; state university; technical institute; professional training
	Senior/adolescent theme
1. listens to	Frank Sinatra; Bing Crosby; Eminem; BB King; Paul Simon; Miles Davis; Stevie Wonder
2. favorite breakfast	oatmeal; cream of wheat; cold pizza; bagel; eggs; muffin; bacon
3. favorite hobby	checkers; genealogy; video games; woodworking; hiking; fishing; photography
4. favorite TV show	*Murder, She Wrote*; *Matlock*; *MTV Real World*; *ER*; *Friends*; *Law and Order*; *Cops*
5. wears	cardigan sweaters; polyester trousers; hooded sweatshirts; business suits; cotton slacks; khaki pants; linen shirts
6. occupation	retired; crossing guard; student; executive; teacher; manager; salesperson
7. dream car	Cadillac; Lincoln Town Car; Ferrari; Volvo; BMW; Taurus; Camry
8. drink of choice	prune juice; decaf tea; Dr. Pepper; coffee; Calistoga water; apple cider; wine
9. health	poor, often ill; very poor, chronically ill; excellent, never ill; average, occasionally ill; moderate, rarely ill; variable, usually good; normal, mild allergies
10. member of	Gray Panthers; Shriners; local skateboarders club; Sierra Club; volunteer fire department; company softball team; neighborhood council
11. favorite movie star is	John Wayne; Fred Astaire; Lindsay Lohan; Meryl Streep; Robert De Niro; Sean Penn; Mel Gibson
12. goes to	shuffleboard games; bingo games; "rave" parties; movies; live theater; standup comedy; auto shows

(*Appendix continues*)

Appendix (*continued*)

Dimension	Values
	Female/male theme
1. favorite sport	boxing; football; figure skating; skiing; swimming; soccer; volleyball
2. hobby	fixing cars; drag racing; shopping; tennis; painting; squash; gardening
3. favorite TV show	*Monday Night Football*; *Fear Factor*; *Martha Stewart*; *Seinfeld*; *Friends*; *The O.C.*; *West Wing*
4. favorite movie	*The Terminator*; *Black Hawk Down*; *Little Women*; *Henry the Fifth*; *Rear Window*; *The Graduate*; *On the Waterfront*
5. favorite books	adventure novels; heroic fantasy; romance novels; biographies of famous people; humorous or satirical; mystery novels; self-help manuals
6. typical clothing accessory	tie; suspenders; purse; jacket; vest; hat; sweater
7. employed as	an engineer; a plumber; a model; a teacher; a realtor; a store clerk; an office manager
8. dream vacation	Super Bowl road trip; big-game hunting; Paris fashion tour; Disneyland; New Orleans; Caribbean beachfront; European holiday
9. favorite store	Piersons Hardware; Big 5 Sporting Goods; Greystone Jewelers; Northtown Books; Sjaak's Organic Chocolates; Costco Wholesale; Sears Retail
10. perfect gift	power tools; guns; flowers; clothes; money; gift certificate; music CDs
11. reads first	sports section; business pages; health and beauty section; front page; classified ads; editorials; local news
12. favorite magazine	*Popular Mechanics*; *Outdoor Life*; *Cosmopolitan*; *Newsweek*; *Rolling Stone*; *Travel and Leisure*; *Wine Spectator*

Note. Stimulus materials (attribute dimensions and values) for the experiments described in the article. Note that seven values are listed for each dimension. The values are listed in the order assumed by the numerical codes used in the Method sections above. Thus, value 1 for the favorite drink dimension would be sherry, value 2 would be fine wine, and so on.

Received March 11, 2011
Revision received January 13, 2012
Accepted February 20, 2012 ■