Memory & Cognition
2002, 30 (1), 107-118

Effect of theory-based feature correlations on typicality judgments

WOO-KYOUNG AHN, JESSECAE K. MARSH, and CHRISTIAN C. LUHMANN
Vanderbilt University, Nashville, Tennessee

and

KEVIN LEE
Yale University, New Haven, Connecticut

In the present study, we examine what types of feature correlations are salient in our conceptual representations. It was hypothesized that of all possible feature pairs, those that are explicitly recognized as correlated (i.e., explicit pairs) and affect typicality judgments are the ones that are more likely theory based than are those that are not explicitly recognized (i.e., implicit pairs). Real-world categories and their properties, taken from Malt and Smith (1984), were examined. We found that explicit pairs had a greater number of asymmetric dependency relations (i.e., one feature depends on the other feature, but not vice versa) and stronger dependency relations than did implicit pairs, which were statistically correlated in the environment but were not recognized as such. In addition, people more often provided specific relation labels for explicit pairs than for implicit pairs; these labels were most often causal relations. Finally, typicality judgments were more affected when explicit correlations were broken than when implicit correlations were broken. It is concluded that in natural categories, feature correlations that are explicitly represented and affect typicality judgments are the ones about which people have theories.

It is generally agreed that our concepts are formed around clusters of correlated features (Rosch, 1978). For instance, *having wings* is more likely to occur with *being able to fly* and *having feathers* than with *being able to swim* and *having gills*. Thus, we give the name *bird* to creatures that have wings and feathers and are able to fly. Within these correlation clusters, there are specific feature correlations that are more important and explicit in our conceptual representations. In our *bird* concept, *having wings* and *being able to fly* would be more closely tied to each other than *having feathers* and *chirping*. The goal of the present study is to examine which feature correlations, among all possible pairs in a cluster, are explicitly represented in concepts and affect typicality ratings of exemplars beyond the contribution of their individual features.

The work on people's naive theories of natural categories (Carey, 1985; Keil, 1989; Murphy & Medin, 1985) sheds some light on our goals. According to this approach, people do not simply notice and record statistically correlated properties when learning categories. Instead, people's knowledge of a domain emphasizes only a subset of the statistical correlations that occur. We believe that this possibility has yet to be tested, especially with natural categories, as will be elaborated upon below. In our studies we will strive to (1) demonstrate that explicitly represented feature correlations in concepts are theory based, (2) examine the content of these naive theories, and (3) demonstrate that these theory-based, explicitly represented feature correlations are more determinative of category members' typicality than are theory-neutral feature correlations. To achieve these goals, we must first review previous studies on the effect of correlated features, particularly noting studies that involved correlated features in real-life categories.

Effect of Correlated Features Found in Previous Studies

Are people sensitive to all correlations among features when learning novel categories? In an attempt to answer this, Medin, Altom, Edelson, and Freko (1982) had participants learn artificial categories. When asked to judge category membership for new items, those that preserved the study phase correlation were more likely to be selected as members, even when the item with the preserved correlation contained fewer typical features (see also Wattenmaker, 1991, 1993). However, these results might not generalize to real-world cases, because the correlation between features was perfect, unlike in natural concepts (e.g., *having wings* and *being able to fly*). Furthermore, in these studies, a rule involving the correlated features was more diagnostic and

This project was supported by National Institute of Mental Health Grant RO1 MH57737, awarded to the first author. We thank Kathy Johnson, Barbara Malt, Thomas Palmeri, and Steven Sloman for their extremely helpful comments on earlier drafts of this paper. We also thank Barbara Malt for providing the list of basic properties and the family resemblance weightings of properties used in Experiment 4. Correspondence concerning this article should be sent to W.-K. Ahn, Department of Psychology, Vanderbilt University, Wilson Hall, 111 21st Avenue South, Nashville, TN 37203 (e-mail: woo-kyoung.ahn@vander bilt.edu).

—Accepted by previous editorial team.

 Copyright 2002 Psychonomic Society, Inc.

simplistic than any rule involving individual features alone, making feature correlations unfairly advantaged over individual features (see also Wattenmaker, 1993). Therefore, it is difficult to conclude that people would be sensitive to *all* correlations among features in natural concepts, as these studies have demonstrated with artificial stimuli.

Using natural categories, Malt and Smith (1984) found that people are sensitive to feature correlations to some degree. Although this work has been regularly cited as a classic study demonstrating sensitivity to correlated features, a close examination reveals that their results are not so straightforward. In their study, about two thirds of the statistically correlated properties in natural categories were not explicitly noticed as being correlated. Furthermore, Malt and Smith did not find that these statistical correlations invariably determined goodness of exemplars in a category. Since the present study is based on the stimuli and procedure used in their study, we will give a detailed description.

Malt and Smith (1984) began by asking what features are statistically correlated within a category. In Part 1 of their Experiment 1, participants were asked to list properties that are generally true of members (e.g., robin, sofa) of various categories (e.g., birds, furniture). Properties that were listed by at least a third of the participants and present in at least two members in a category were then provided to another group of participants, who rated the properties as to how much each applied to each category member. An exemplar × property table was created, containing mean ratings for every category member on all properties. Pearson correlations were conducted on the mean ratings across category members for all possible pairs of properties. For instance, if two properties are correlated in a category, category members with high mean ratings on one should have also received high ratings on the other. On average, 33.5% of the property pairs were significantly correlated within a category at the .05 level. Henceforth, we will call these pairs *statistically correlated feature pairs*.

The next question Malt and Smith (1984) raised was whether these statistically significant correlations affect typicality judgments of exemplars. In Part 2 of their Experiment 1, participants received exemplars in which these correlations were either preserved or broken, while family resemblance scores[1] were equated to examine the correlations' effects on typicality judgments separately. Surprisingly, they failed to find an effect. That is, unlike in the previous studies using artificial stimuli (e.g., Medin et al., 1982), people were not sensitive to the correlated structures presented in natural categories.

One explanation that Malt and Smith (1984) offered for this failure was that property correlations were statistical co-occurrences and were not necessarily present in people's representations. In their Experiment 2, participants rated the extent to which they thought each pair of properties was related within the category. Of the original pairs, only 33% were explicitly rated as correlated. Therefore, people were not aware of all feature correlations present in a category. In Part 2 of their Experiment 2, they used only correlated features in their exemplars that were explicitly judged to be correlated. Participants now found exemplars with intact correlations to be more typical than exemplars breaking such correlations.

The specific question investigated in our present experiments is why only a subset of the statistically significant correlations were recognized as being correlated and why only this subset affected typicality judgments. Malt and Smith (1984) end their article by pointing out the need for this very investigation:

> Exactly what that subset of [correlated] pairs consists of, however, has not been systematically explored. Whether they are primarily pairs that are functionally related—such as large birds requiring large wings in order to fly—or are empirically plausible combinations—such as birds far from the ocean being unlikely to find fish to eat—or are simply the most frequently occurring combinations, is not clear. One step toward better understanding the human categorization process would seem to be in identifying what types of properties and property pairs are used. (p. 269)

The Role of Background Knowledge

The number of possible correlations in the world is disturbingly large (e.g., Keil, 1981; Murphy & Medin, 1985). For instance, in a category with only seven binary-valued dimensions, 84 correlations exist. Out of these, which types would people notice?

One way of getting around the computational complexity is to be guided by existing background knowledge. Numerous studies have demonstrated how people's judgments of correlation between two events are influenced by their initial hypothesis (e.g., Chapman & Chapman, 1967; see Alloy & Tabachnik, 1984, for a review). These findings should also hold in detecting correlations within concepts. If people understand the reasons behind correlations, these correlations will be noticed more easily. For instance, consider some of the statistical correlations found in Experiment 1 of Malt and Smith (1984): *Being gray/white* in birds was positively correlated with *eating fish*, and *having buttons* and *having long sleeves* were positively correlated in clothing. Yet, the causal links underlying these correlations are not immediately apparent. In contrast, most laypeople would know causal connections underlying the correlation between *being made of wool* and *being warm* for clothing, or *living near the ocean* and *eating fish* for birds. Indeed, the latter pairs, and not the former, are the ones that were shown to have been explicitly noticed in Malt and Smith. Thus, we propose that people tend to notice correlations in concepts if they are meaningful to them in light of their existing domain theories.[2]

We further suggest that it is explainable correlations that influence typicality judgments. According to the *theory-based* approach to categorization, concepts are embedded in theories and are represented by causal links and explanatory relations. Murphy and Medin (1985) state that "the connection between those [correlated] features is not a simple link, but a whole causal explanation for how the two are related" (p. 300). From this perspective, conceptual coherence is based on explanations or domain theories,[3] rather than on clusters of undifferentiated correlations. If people view conceptual coherence this way, exemplars that preserve theory-based correlations will be judged to be better category members than are exemplars that do not preserve such correlations, all else being equal. On the other hand, mere statistical correlations that are not linked by explanatory relations would not influence membership judgment as much, because these relations are not critical in conceptual representations. Therefore, our second hypothesis is that the types of correlations that should matter in typicality judgments are the ones that are linked within one's domain theories.

Unfortunately, few previous studies have directly examined the relationship between background knowledge and sensitivity to correlations in natural categories. Some studies used artificial stimuli to examine this issue, but they were limited in various ways. For instance, Barrett, Abdi, Murphy, and Gallagher (1993) found that children were more sensitive to theory-based correlations (e.g., *having a big brain* and *a good memory*) than to theory-neutral correlations. As the authors acknowledged, however, this result could have been due to a simple heuristic (e.g., big things are good, so big brains go with good memory), rather than to children's specific domain theories. In their Experiment 2, designed to avoid this problem, children had more difficulty; more than 50% of the participants were dropped for failing to show consistent responses. Thus, it is difficult to tell how prevalent the effects of theory-based correlations really were.

Murphy and Wisniewski (1989) studied the effect of theory-based correlations in adults' categorization of natural concepts. In this study, when a feature in a theory-related pair (e.g., *eats fish* and *lives under water*) was substituted with a feature in a contrast category (e.g., *eats wheat* and *lives under water*), the broken relation significantly lowered category membership judgments. As in this example, broken correlations in this study resulted in bizarre examples, violating layperson's ontological assumptions (e.g., Keil, 1981). Murphy and Wisniewski proposed that people use their domain theories to rule out implausible or bizarre co-occurrences of features within a concept. Although this certainly would be one mechanism through which prior knowledge can play a role in concept learning, it does not answer the initial question that we posed in relation to Malt and Smith's (1984) results. In Malt and Smith's study, the effect of correlated features was found even when the items with broken correlations were not bizarre.

To summarize, previous studies have shown that background knowledge affects people's sensitivity to correlations and suggested possible underlying mechanisms (e.g., background knowledge rules out implausible correlations). However, none of the studies has provided strong evidence that the presence of naive domain theories accounts for why certain correlations mattered more in typicality judgments of real-life concepts, as was found by Malt and Smith (1984).

Overview of Experiments

We carried out four experiments to test the hypothesis that domain theories determine which feature correlations would be salient in our representation of natural categories. First, in Experiment 1, we determined which feature correlations were salient in people's conceptual representations by replicating Malt and Smith's (1984) study. In Experiments 2 and 3, we examined whether the feature correlations that were salient were more theory based than the ones that were not salient. Finally, in Experiment 4, we tested whether these two types of feature correlations would differentially affect typicality judgments of category exemplars.

EXPERIMENT 1

In Experiment 1, we examined which feature correlations in concepts people explicitly notice. This is a crucial step, because the explicitness of relations serves as the main independent variable for the remaining experiments. Although Malt and Smith (1984) already had provided such data, we replicated this study with the same participant population as in the rest of the experiments.

Method

Participants. Twenty-one undergraduates from Vanderbilt University participated in partial fulfillment of introductory psychology course requirements. One of these participants was excluded from the study because more than 50% of his/her responses were made within 500 msec, which was deemed not enough time to read the stimuli under normal circumstances.

Materials and Procedure. The procedure was modeled after Part 1 of Malt and Smith's (1984) Experiment 2, with modifications being made to support a computer presentation of the stimuli. We used property pairs found to be statistically correlated in Malt and Smith's Experiment 1 for six different categories. These categories were Bird, Clothing, Flower, Fruit, Furniture, and Tree. The participants were shown one property pair on the computer screen at a time, with the appropriate category name appearing at the top of the screen in capital letters. The participants were told that for each pair, they were to "make a judgment as to how well the features go together in that category." The participants were instructed to use a 7-point scale ranging from −3 (*a strong negative relation*) to +3 (*a strong positive relation*), with a midpoint of zero standing for *no relationship*. The participants were given examples of feature pairs within the category *Deer* that would fit both of the scale's extremes (*is male* and *has horns* for the positive extreme and *is female* and *has horns* for the negative extreme) and the midpoint (*has horns* and *lives near a lake*) during the initial instructions. During pair presentation, a reminder scale that depicted the rating range was shown at the bottom of the computer screen. The pairs were blocked by category, and pair presentation was randomized within each category. Each subject received all 178 correlated property pairs. The experiment was programmed using SuperLab, Version 1.75.

Results and Discussion

The goal of this experiment was to obtain updated data on what feature pairs were explicitly noted as correlated. For each feature pair, a mean rating was computed across all participants. Following the criterion used in Malt and Smith (1984), pairs rated as greater than or equal to +1.5 for positive correlations and less than or equal to −1.5 for negative correlations were defined as saliently related pairs. These pairs will from now on be referred to as *explicit pairs*. Feature pairs that fell within the cutoffs will be referred to as *implicit pairs*.[4] Overall, 51 feature pairs were explicitly rated as correlated, representing 29% of the total pairs. The percentage of pairs explicitly rated as correlated varied between categories, ranging from 13% for the clothing category to 44% for furniture. A list of the explicit pairs and their mean ratings can be found in Appendix A.

Our results were fairly consistent with Malt and Smith's (1984) results, in that 83.1% of the pairs were consistent in their classification across the two studies. Of the 30 pairs that showed discrepancy, 19 were within a ±0.5 range from a cutoff point, indicating that most of the discrepancy was due to marginal cases. Unless noted otherwise, we use the results from Experiment 1 in the subsequent experiments when we make a distinction between explicit and implicit feature pairs.

EXPERIMENT 2

Experiments 2 and 3 test the hypothesis that in real-world concepts, the kinds of correlations that people explicitly notice are the ones for which people have explanations. In Experiment 2, we treat theory-related correlations in a relatively content-free manner by assuming that they come in only one kind (Thagard, 1989). Following Sloman, Love, and Ahn (1998), we call these unlabeled relations *dependencies*. The term is general enough to cover causal, categorical, temporal, or indeed any other kind of directional relation between features. For instance, a piece of furniture is comfortable because it has cushions, and therefore, *being comfortable* depends on *having cushions*. In contrast, people might not believe that *being gray/white* in birds depends on *living near the ocean* or vice versa. Asking participants only about dependency relations (e.g., "Does *being able to fly* depend on *having wings*?") is a less confusing and more succinct procedure than listing all possible theory-based relations in a question (e.g., "Does *having wings* cause/allow/follow/is depended upon by/trigger/increase/decrease/etc. *being able to fly*?").

Experiment 2 used pairs of features that were found to be statistically correlated in Malt and Smith's (1984) Experiment 1. As was found in our Experiment 1, some of these pairs were more explicitly noted as being correlated (explicit pairs) than were others (implicit pairs). For each pair of features (say A and B), we asked participants to rate the strengths in two directions (i.e., the degree to which A depends on B, and the degree to which B depends on A). If a person believes that A causes/allows/is followed by/triggers/increases/decreases/etc. B, the rating on "B depends on A" would be high, whereas the rating on "A depends on B" would be low. We predicted that this would be a dominant pattern in explicit pairs, because we hypothesized that people have specific explanations for the correlations that they explicitly notice. On the other hand, if a person represents the pair as being merely correlated, without any explanations for why, both ratings will be low. For instance, *having four legs* and *having cushions* for furniture were found to be statistically correlated, but *having four legs* does not depend on *having cushions*, nor does *having cushions* depend on *having four legs*. We predicted that this would be a dominant pattern for the implicit pairs. To summarize, we predicted that in the explicit pairs, one of the two dependency judgments within a pair would be high, whereas in the implicit pairs, both would be low. Thus, a maximum dependency rating within a pair would be higher in the explicit pairs than in the implicit pairs. In addition, the difference between the two dependency judgments within a pair would be higher in the explicit pairs (because B depends on A but A does not depend on B) than in the implicit pairs (because both judgments would be low in the implicit pairs).[5]

Method

Participants. Eighteen undergraduates from Yale University and 10 undergraduates from Vanderbilt University participated in this experiment in partial fulfillment of an introductory psychology course credit.

Materials. The same property pairs as those from Malt and Smith (1984) were used, as in our Experiment 1. For each property pair, two statements were generated, one to reflect each direction of possible dependency. For example, using the properties *eats fish* and *lives near the ocean*, two test items were generated: "For a bird, whether or not it eats fish depends on whether or not it lives near the ocean" and "For a bird, whether or not it lives near the ocean depends on whether or not it eats fish." Since there were 178 feature pairs in our study, this led to a total of 356 statements (106 explicit statements and 250 implicit statements). Because responding to all 356 statements is too demanding, each participant received roughly half of these statements. Data from our Experiment 1 was not available at the time of this study, so we used Malt and Smith's distinction of explicit and implicit feature pairs in sampling statements for each participant. Using Malt and Smith's distinction, all the participants received all of the explicit statements, but they received approximately half of the implicit statements, sampled across six categories.[6] In this way, the number of implicit and explicit statements for each participant was equated as much as possible. Note that Malt and Smith's explicit/implicit distinctions were used only when assigning the items to the participants. All the analyses reported in this study are based on the explicit/implicit distinctions obtained from our Experiment 1.

Procedure. The participants were instructed to judge the degree to which they agreed with each statement on a 9-point scale anchored with *strongly disagree* for 1 and *strongly agree* for 9. The participants were alerted to be sensitive to asymmetric dependency relations, rather than simple co-occurrences, as follows:

> Also remember that some relationships might be plausible in one direction, such as: "For a chair, whether or not you can sit on it depends on whether or not it is sturdy," but not in the other direction: "For a chair, whether or not it is sturdy depends on whether or not you can sit on it." Be sure to pay attention to which property is depended upon (the second property in the statement) and which is dependent (the first property).

After reading the instructions, the participants proceeded to rate randomly presented statements. The experiment was programmed using Psyscope 1.1 (Cohen, MacWhinney, Flatt, & Provost, 1993).

Results

Do explicit pairs have stronger dependency relations than do implicit pairs? To answer this question, we examined what we termed *maximum dependency scores* within each pair. Recall that the participants rated dependency strengths in two different directions within each pair (A to B and B to A). The larger of the two ratings within each pair (i.e., the maximum dependency score) is the dependency strength of that pair. For each participant, we obtained an average maximum score for the explicit pairs and an average maximum score for the implicit pairs. A paired t test showed that maximum scores were significantly higher for the explicit pairs (6.18) than for the implicit pairs [3.62; $t(27) = 15.48$, $p < .0001$]. As is shown in Table 1, the direction of difference was consistent across all six categories. Therefore, the results strongly support the conclusion that explicit pairs have stronger dependency relations than do implicit pairs.

We also examined the difference score within each pair. As was explained earlier, if, for instance, A causes B, then the difference between the rating on "A depends on B" and "B depends on A" would be large. However, if A merely correlates with B, the difference would be small.[7] For a statistical analysis, we first calculated a difference between a rating on "A depends on B" and a rating on "B depends on A" in each pair for each participant. The average difference scores for the implicit pairs and those for the explicit pairs within each participant were used as input for a paired t test. The results showed that difference scores were significantly higher for explicit pairs (2.09) than for implicit pairs [1.47; $t(27) = 5.04$, $p < .0001$]. As is shown in Table 1, the results were consistent across all six categories.

Discussion

To summarize, we found that explicit pairs contained stronger dependency links than did implicit pairs, as was shown by the difference between the two kinds of pairs with respect to maximum scores. Furthermore, explicit pairs had more asymmetric dependency relations. These results support the hypothesis that explicitly noticed correlated pairs are more theory based than are those that are not explicitly noticed.

Relying only on dependency relations had an advantage in that the participants were not concerned about how to classify or label relations, making the task easier. However, this method might have been limited in determining which feature pairs are theory based; dependency relations might have been either too broad or too narrow as operational definitions for theory-based relations. Each will be discussed.

First, dependency relations might have been too broad a measure in that they might have also captured nontheoretical, asymmetric relations. For instance, most laypeople would be able to tell that in tires P(made by Goodyear | black) is lower than P(black | made by Goodyear). Thus, the relationship between the two is asymmetric, but it is very unlikely that people know any reason for this relationship. We highly suspect that in a case like this, people would not have given a high dependency rating. That is, it is quite unlikely that people would say that "whether or not a tire is made by Goodyear depends on whether or not a tire is black" or "whether or not a tire is black depends on whether or not a tire is made by Goodyear." The phrase *depends on* implies much more than just a conditional probability. Therefore, both the maximum and the difference dependency rating would have been lower in this kind of nontheoretical, asymmetric case. Yet, it is still possible that dependency relations might have implied nontheoretical, asymmetric relations, and the difference between the explicit and the implicit pairs might have been due to these nontheoretical ones. Experiment 3 will address this issue.

Alternatively, using dependency as an umbrella for all theory-based relations might have been a gross simplification; there may have been many other theory-based relations that are not easily captured by the term *depends on*. This, however, does not necessarily pose a problem for Experiment 2, because according to this argument, the use of dependency relations would have underestimated the differences between explicit and implicit pairs. Yet, this could still be a problem if most of the implicit pairs happened to consist of theoretical relations that could not have been described as dependency. In this case, we cannot conclude that the explicit pairs are more theory based than are the implicit pairs, despite the fact that the explicit pairs had stronger dependency relations than did the implicit pairs. Although this possibility would be a large coincidence, Experiment 3 was carried out to also address this concern.

Table 1
Mean Maximum and Difference Scores for Explicit and Implicit Pairs

	Maximum Scores		Difference Scores	
Category	Explicit	Implicit	Explicit	Implicit
Bird	6.30	3.21	1.95	1.16
Clothing	5.63	3.38	2.59	1.48
Flower	7.13	2.15	1.90	0.90
Fruit	6.30	3.78	2.48	1.52
Furniture	5.76	4.30	2.06	1.77
Tree	6.28	3.71	2.47	1.46
Average	6.18	3.62	2.09	1.47
SD	1.26	1.09	0.54	0.50

EXPERIMENT 3

Experiment 3 uses a more open-ended task by letting participants draw relations by using any labels that they want. The participants were presented with pairs of features and were asked to draw an arrow with a label on top to indicate any specific relation that they believe existed. They were specifically told not to draw mere correlations. Thus, it was possible to examine whether feature correlations that are explicitly noticed were indeed based on specific theory-based relations, rather than on merely salient correlations.

From these data, we also examined the content of theory-based relations. Proponents of theory-based categorization (e.g., Carey, 1985; Murphy & Medin, 1985; Wellman, 1990) have emphasized causality as a central component in conceptual or theory representations. We examined whether this indeed was the case.

Method

Participants. Twenty-one Vanderbilt University undergraduates participated in partial fulfillment of the requirements of an introductory psychology course. One participant did not finish the task and was therefore excluded from the study.

Materials. Again, the property pairs found to be statistically correlated by Malt and Smith (1984) were used. Pairs were arranged on

Figure 1. Sample item from Experiment 3. The label in italics and the arrow represent a hypothetical addition by a participant.

paper in a table format. Each row of a table was divided into two columns: a large column on the left and a small column on the right. Within the left column, a feature pair was presented with a large gap between the features. The right column was left empty. Figure 1 represents what each item looked like after a participant added a labeled arrow and rated the feature pair.

Property pairs were grouped by category, with the category name appearing at the top of each page. Within each category, the pair order was randomized so that no property was presented in more than two consecutive pairs. A second version of the packet was created by reversing the order of the properties within pairs (i.e., a feature that appeared on the left in a pair in Version 1 appeared on the right in Version 2, and vice versa) and rerandomizing the property pairs within categories. Ten participants received each version. For each participant, the order of the six categories within the packet was randomized. Each subject received all 178 property pairs.

Procedure. The participants were asked to decide whether a relationship existed between the members of a feature pair. If they believed that there was a relationship between the members of a feature pair, they were asked to (1) draw an arrow indicating the directionality of the relationship, (2) label the arrow as to the type of relationship it represented, and (3) rate the strength of the relationship on a scale of 1 to 5, where 1 represented a *weak relationship* and 5 represented a *strong relationship*. The participants were given the following list of suggested labels: depends on, precedes, is a precondition for, allows, enables, affects, causes, is a subset of, is an example of, increases, and decreases. This list was created by the first three authors of this article and was intended to be exhaustive. To include any plausible relations that were not thought of, the participants were encouraged to use labels of their own invention where they felt it to be appropriate. The participants were instructed that if they believed a bidirectional relationship existed between members of a pair, to draw arrows in both directions and label both accordingly. The participants did all arrow drawing and labeling in the space between features in the table and recorded the strength of the relationship in the empty right column. If the participants did not believe a relationship existed between members of a feature pair, they were instructed to leave the area between a feature pair blank and to write a zero in the corresponding strength box.

The participants were told that any feature pair that they were shown co-occurred within the given category and, for this reason, to refrain from labeling relationships as *co-occurs* or *associated*. They were encouraged to use more precise labels that reflected a more specific relationship. If they could not think of a more specific label, they were asked to rate the pair as *no relationship*.

Results and Discussion

First, we looked at what percentages of explicit and implicit pairs were given a specific relationship label by subjects. The results showed that explicit pairs were much more likely to have been given a specific relation (68.9%) than were implicit pairs (43.4%). A paired t test was performed on proportions of labeled relations in explicit and implicit pairs for each participant, and a significant difference was found between the two types of property pairs [$t(19) = 8.45$, $p < .001$]. Another paired t test was carried out on mean strength ratings for explicit and implicit pairs for each participant (with no relations treated as having a strength of 0), and it was found that explicit pairs (2.50) had significantly stronger relations than did implicit pairs [1.33; $t(19) = 12.16$, $p < .0001$]. As is shown in Table 2, the direction of difference was consistent across all six categories.

Yet, one might be concerned that nearly half the time (43.4%), the implicit pairs were labeled to have specific relations. However, the implicit pairs that fell close to the designated cutoff of ±1.5 made up the majority of the implicit pairs found to have a specific relation. Indeed, the correlation between the mean number of specific relations and the mean explicitness ratings from Experiment 1 was significantly positive even within implicit pairs only ($r = .57$, $p < .001$).

Taken together, the results strongly support the claim that the more explicit a pair is, the more likely people know specific relations. In other words, when people explicitly notice correlations of features in concepts, they do not simply code them as mere correlations; they have beliefs about what these relations are.

In addition, a correlational analysis was carried out between the strengths of the specific relations found in Experiment 3 and the strengths of the general relations found in Experiment 2, and a significantly positive correlation was found ($r = .71$, $p < .01$). This result indicates that the measure of dependency strengths generally converge with the measure of specific relations used in Experiment 3.

Then what is the content of these relations? Some of the labels that participants produced could be easily grouped as the same kind. For instance, as is shown as notes of Appendix A, *decreases* and *minimizes* are grouped as the same label. Through discussion, the first three authors of the present article grouped the labels into seven categories. Appendix A explains the grouping and lists the relations that were mentioned by at least 4 participants for each explicit pair, with the number of participants who mentioned that label in parentheses. As can be easily seen, causal relations were most frequently listed (52.4% of all labeled relations).

In order to provide more objective, quantifiable indices, we measured how *causal* each relation was from an independent group of 8 volunteers (4 undergraduates and 4 re-

Table 2
Mean Strengths of Relations for Explicit and Implicit Feature Pairs in Experiment 3

	Type of Feature Pairs	
Category	Explicit	Implicit
Bird	2.51	1.03
Clothing	3.69	1.66
Flower	2.54	1.25
Fruit	3.23	1.78
Furniture	2.33	1.83
Tree	3.20	1.29

search assistants with bachelor's degrees). They received a list of all the labels produced by the participants of Experiment 3. Then they rated each label on "the extent to which the term implies that there is a (positive or negative) causal link underlying the two items" on a 9-point scale anchored with *term definitely does not imply a causal link* for 1 and *term definitely implies a causal link* for 9. A mean rating for each label was obtained and used as a *causality index* indicating how causal that label is. These indices ranged from 2.38 ($SD = 1.99$) for *is an example of* to 8.88 ($SD = 0.35$) for *causes*. Then, for each feature pair used in Experiment 3, we calculated an average causality index of all relations mentioned for that pair. For instance, suppose 10 out of 20 participants listed *causes* as a relation for a feature pair, and the other 10 did not list any relation. The average causality index for that feature pair would be 8.88. Thus, for each pair, we came up with a measure of how causally the two features are related, given that it is judged as having a specific relationship. Across all feature pairs, the mean causality index was 5.67, which was significantly higher than 5 [i.e., a midpoint on the 9-point scale used to measure causality indices; $t(176) = 10.37$, $p < .001$]. This suggests that, overall, the specific relations people used to explain feature correlations are more causal than noncausal. Interestingly, explicit pairs were judged to be more causal ($M = 6.18$) than were implicit pairs [$M = 5.46$; $t(175) = 5.49$, $p < .001$].

To summarize, we started out with a question of why some feature correlations are more noticeable than others when all are statistically significantly correlated. In Experiment 3, we found that feature correlations that people tend to be aware of differ from feature correlations that they tend to be unaware of, in terms of whether people can provide explanations for how they are related. People were more likely to provide specific labels for the explicit pairs than for the implicit pairs. In addition, we found that causality is an important component in these explanations, especially for the ones provided for the explicit pairs.

One final point to be discussed is the finding that a majority of the participants came up with theory-based relations between pairs that had a category name as a feature (i.e., being a bird). As is shown in Appendix A, three such cases were *is a bird* causes/allows, is a precondition for *has feathers*, and *has a beak* and *is a bird* also causes/allows *lays*. One might argue that these relations are more definitional than explanatory: Animals that have feathers and beaks and that lay eggs are called *birds*. Subsequent analyses showed that the pattern of all results reported in this study still holds without these pairs. In addition, we argue, on the basis of psychological essentialism (Medin & Ortony, 1989), that these relations are truly believed to be explanatory. That is, people would have meant that the *bird essence* causes birds to have feathers and a beak and to lay eggs.

EXPERIMENT 4

In Experiments 2 and 3, we found that correlations that are explicitly noticed are theory based. Now, we return to the final question of the impact of preserving these correlations on typicality judgments. Malt and Smith (1984) were interested in *whether* feature correlations could influence typicality judgments. In the end, they could demonstrate the effect when they utilized explicit pairs only. We have taken their findings a step further to see what type of statistical correlations matter most in goodness of exemplars. We hypothesize that explicit (or theory-based, as was found in our experiments) correlations are much more likely to influence typicality judgments than are implicit (or theory-neutral) correlations. In Experiment 4, participants were presented with pairs of items that either broke or preserved a feature correlation and then judged which item was more typical of a target category.

Method

Participants. Thirty-three Vanderbilt University undergraduates participated in partial fulfillment of the requirements of an introductory psychology course. Three of the participants were excluded from the study because they did not correctly indicate their responses in the provided packets.

Materials. From the results obtained in Experiment 1, we selected all of the property pairs that had been rated as positive explicit relations.[8] These pairs were used to construct artificial exemplars of the six categories, using the same method as in Malt and Smith (1984).

Specifically, the exemplars for a given category had three common basic properties. These properties were identical to the ones used in Malt and Smith (1984) and are listed in Appendix B under "Constant Features." In addition, each exemplar had two critical properties that either formed an explicit pair or did not. In the cases in which the explicit pair relation was broken, one of the two properties was replaced by another property of that category that had an equal or greater family resemblance weight as used by Malt and Smith.[9] As in Malt and Smith's study, it was ensured that each replacement property was not rated as explicitly related or negatively related to the remaining property in the results of our Experiment 1. Of the original 35 explicit pairs that were taken under consideration, 25 pairs met the criteria for having matching family resemblance weights and were therefore used to create exemplars. These pairs are noted in Appendix A. The pairs varied in number across categories in the following way: 9 for Furniture, 6 for Tree, 4 each for Bird and Clothing, and 1 each for Flower and Fruit.

After the selection of the 25 explicit pairs, 25 implicit pairs were chosen, using the same selection procedure as that described for the explicit pairs, and with an additional constraint that there be an equal number of implicit pairs and explicit pairs within each category. The implicit pairs were then used to create implicit exemplars of the same format as the explicit exemplars.

In order to verify that the 25 explicit pairs selected for Experiment 4 are more theory based than the 25 implicit pairs, these two sets of pairs were compared with respect to the two measures used in Experiments 2 and 3. The maximum dependency ratings measured in Experiment 2 were significantly higher for the selected explicit pairs ($M = 6.38$) than for the selected implicit pairs [$M = 3.45$; $t(48) = 8.21$, $p < .0001$]. Similarly, the strengths of the explanatory relations measured in Experiment 3 were significantly higher for the selected explicit pairs ($M = 3.22$) than for the selected implicit pairs [$M = 1.11$; $t(48) = 10.90$, $p < .0001$].

Exemplars were created so that each possessed three basic properties and two critical properties, except for one bird and one tree exemplar. These two exemplars had four features, because one of the critical pair features was the same as a constant feature used by Malt and Smith (1984; i.e., *has feathers* and *has leaves*). For each stimulus, the constant properties were listed first, followed by the critical pair.

An exemplar preserving the original explicit or implicit relation and an exemplar breaking this relation appeared on each page of the experimental packet. In half of the explicit and implicit pair items, the preserved exemplars were placed on the left side of a page, and in the other half they were placed on the right side of a page. In all of the items, the position of the feature that differed between the two sides was randomly placed as either last or second to last in the feature list. Exemplars were blocked by category, and category presentation was randomized for each participant. Within each category, explicit and implicit exemplars were randomly mixed by page. A second set of materials was created from the first version by reversing the side on which the preserved correlation could be found. The second set of materials was identical to the first in every other respect. The instructions for the booklet were typed on the first page of the experimental packet.

Procedure. The procedure was identical to Part 2 of Malt and Smith's (1984) Experiment 2. To reiterate, the participants were provided with pairs of exemplars and were asked to choose which of the pair, the left or the right, seemed more typical of the category listed at the top of the page. They were then asked to rate both exemplars on a 1 to 7 scale as to exactly how typical each was of its category, where 1 was *very atypical* and 7 was *very typical*. They were told to avoid making their decisions by simply comparing the two properties that differed and, instead, to consider each exemplar as a whole.

Results and Discussion

As was predicted, the participants were much more likely to choose exemplars preserving a correlation when the correlation was an explicit type (or theory based, as was found in previous experiments) than when it was an implicit type (or theory neutral). For each participant, the mean number of selections of an exemplar with a preserved correlation over a broken correlation was calculated for explicit and implicit pair exemplars. A paired t test based on these scores showed a significant difference between explicit pair exemplars (M = 70.8%) and implicit pair exemplars [M = 49.0%; $t(29) = 9.362$, $p < .0001$]. In fact, choices on implicit pair exemplars did not differ from the chance level of 50% [$t(29) = 0.64$, $p = .53$]. Another t test was performed by treating items as random variables, and a significant difference was found [$t(48) = 3.31$, $p < .01$].

Finally, the participants' typicality ratings were analyzed. For each participant, a mean rating was obtained for each of the four types of feature pairs (i.e., correlation-preserved and correlation-broken items for explicit and implicit feature pairs). The overall means for these four types of feature pairs are shown in Figure 2. A repeated measures analysis of variance found a significant main effect of whether or not correlations were preserved [$F(1,29) = 30.26$, $MS_e = 0.026$, $p < .001$] and a significant main effect of whether correlations are explicit or implicit [$F(1,29) = 7.32$, $MS_e = 0.056$, $p < .05$]. But these main effects should be interpreted in light of the significant interaction between these two factors [$F(1,29) = 16.34$, $MS_e = 0.042$, $p < .001$]. The interaction was significant because the typicality ratings on items with preserved explicit correlations (M = 5.11) were significantly higher than the typicality ratings on items with broken explicit correlations [M = 4.79; $t(29) = 6.24$, $p < .001$], whereas the typicality ratings on items with preserved implicit correlations (4.84) did not differ from those on items with broken implicit correlations [4.83; $t(29) = 0.23$, $p = .82$]. Therefore, the results clearly supported the hypothesis that explicit correlations, but not implicit correlations, matter in typicality judgments.

Another way of looking at this interaction is to compare each of the preserved and broken correlations across the explicit and the implicit pairs. For items with broken correlations, there was no difference between the explicit pairs and the implicit pairs [$t(29) = 0.60$, $p = .55$]. However, the items with preserved explicit correlations were judged to be significantly more typical than the items with preserved implicit correlations [$t(29) = 4.69$, $p < .0001$]. That is, the interaction was not obtained because breaking correlations supported by domain theories greatly lowered typicality ratings. Instead, it was obtained because preserving correlations supported by domain theories increased the typicality of exemplars. These results contrast with Murphy and Wisniewski (1989), who found that the effect of theory-based correlations was obtained because items with broken correlations were bizarre in light of domain theories (e.g., *eats wheat* and *lives under water*), as was discussed earlier in the introduction. That is, according to their proposal, background theories matter because domain theories rule out implausible cases. Our findings show that preserving theory-based correlations could even enhance the typicality of exemplars.

GENERAL DISCUSSION

It has often been described that, in concepts, feature correlations affect categorization beyond the contribution of their individual features. Medin et al. (1982) provides one of the strongest sources of evidence for this. Yet, we argued that their results from artificial stimuli are difficult to generalize to natural concepts. Using feature correlations that were intended to be more similar to those in natural concepts, Murphy and Wisniewski (1989) failed to find any evidence that people base their categorization decisions on the statistical relation of individual features. When nat-

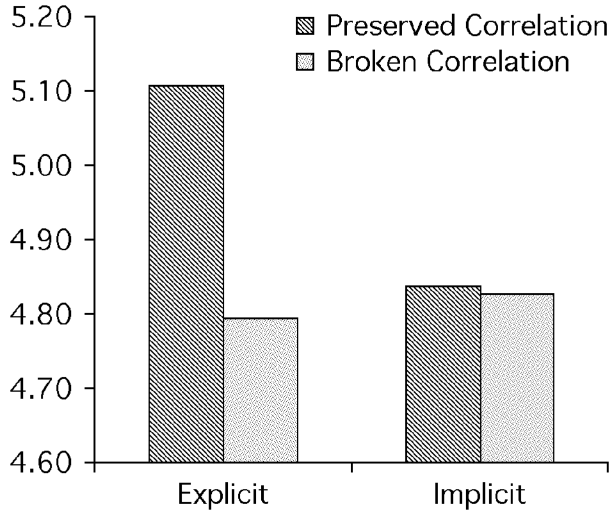

Figure 2. Mean typicality ratings for items that preserved or broke explicit or implicit correlations.

ural categories are examined, Malt and Smith (1984) found that only about one third of the feature correlations that were statistically significant were noticed as being correlated. Furthermore, Malt and Smith failed to find any effect of feature correlations when all statistically significant feature correlations were examined. Instead, feature correlations that were explicitly noticed as being correlated influenced typicality ratings over and above each individual feature's family resemblance score. Our main goal was to examine why only a subset of feature correlations are explicitly noticed and influence categorization.

The present study provides evidence suggesting that the correlations that are present in our representations of real-world concepts are the ones with explanatory relations. In Experiment 1, we found that people did not explicitly recognize the relationships between all the features that statistically co-occurred within a category, replicating Malt and Smith's (1984) results. In Experiment 2, we found that the pairs people explicitly recognized had stronger dependency relations. In Experiment 3, we found that explicitly recognized feature correlations were much more likely to be associated with specifically labeled relations and that those relations tended to be associated with causality. In Experiment 4, we demonstrated that preserving these explicit, theory-based correlations mattered in typicality ratings. Thus, an item with a theoretically related pair (e.g., *flies* and *sits in trees* for birds) was judged to be a better member of a category than was an item that did not preserve such a correlation (e.g., *has a beak* and *sits in trees*), even when individual features in both items were equally rated as belonging to that category. In contrast, no such effect of feature correlations was observed with implicit feature correlations that the participants failed to notice in Experiment 1 and were less likely to label with specific relations as found in Experiment 3.

The present study has important implications for both the theory-based approach and the similarity-based approach to categorization. With respect to the similarity-based approach, the present study offers an important constraint. For instance, Gluck and Bower (1988; see also Gluck, 1991) describe a configural-cue network model of human learning that represents stimulus patterns in terms of their individual features and pairwise conjunctions of features. However, as was noted earlier, a problem with this approach is computational complexity; there are too many conjunctions of features to keep track of as the number of features increases (Medin, 1983). The present results suggest that, as a way to avoid this complexity, people may tend to explicitly encode only those feature conjunctions that they can explain. That is, they do not encode all conjunctions of features in a content-free manner, as in the configural-cue network model.

These findings also contribute to the theory-based approach by allowing us to examine the effect of theory-based correlations in real-world concepts. Although it is often difficult to analyze the complex network of properties and interrelations found in naturally occurring concepts, we have demonstrated that the influence of background knowledge (e.g., theories) does indeed extend beyond the confines of artificially created stimuli. Thus, we found that theoretical knowledge has a significant influence on the perception of coherence in real-life categories. As was mentioned above, it has been shown that one's domain theory can help rule out incoherent exemplars (Murphy & Wisniewski, 1989). In the present study, we have shown that explanations underlying feature correlations can increase the typicality of an item above and beyond the impact of individual feature frequencies, whereas mere statistical correlations that could not be explained failed to do so. This finding provides evidence of a robust relationship between conceptual coherence and explanatory coherence. Finally, we have provided the first empirical evidence for the general assumption that causality constitutes a majority of relations in laypeople's representations of natural concepts (Carey, 1985; Wellman, 1990).

Our results focused on the effect of theory-based correlations only on typicality judgments. Will theory-based correlations influence other kinds of behaviors associated with concepts, such as inductive inferences? We speculate that the answer is yes. Lassaline (1996) found that inductive strengths of features are greater when the features are bound by causal relations than by noncausal relations. For instance, participants were given the descriptions of two animals (e.g., Animal A and Animal B). They were told that each animal had a property (e.g., Property X) and that, additionally, Animal B possessed another property (e.g., Property Y). Furthermore, the participants were told that, for Animal B, there was a relationship between X and Y. The participants then had to judge the likelihood that Animal A also possessed Y. The relationship between X and Y was described as causal in her Experiment 2 and as noncausal (a temporal relation was used instead) in her Experiment 3. Only when the causal description was supplied did the inductive strength increase beyond the strength generated by the other information alone. On the basis of this, it is reasonable to expect that correlations not supported by any causal relations would not increase inductive strengths as much. Other works have also demonstrated that people's causal background knowledge determines how important a feature is in categorization (Ahn, Kim, Lassaline, & Dennis, 2000), as well as in similarity judgments (Ahn & Dennis, 2001). Taken together, these findings suggest that explanatory relations, especially causal ones, play a crucial role in various aspects of our concepts.

The results of the above-mentioned studies lead to a pressing question for future study: Do people explicitly notice correlations because they can explain them, or do people impose explanations after they explicitly notice correlations? For example, people might explicitly notice a correlation between *being a tire* and *black color* because the correlation is unusually strong or because somebody pointed out the correlation to them, but initially, they have no structure binding those two properties other than a perceived correlation. Later, they might develop some theories as to why tires have to be black (e.g., it is less likely to show dirt). On the other hand, it may be that feature correlations are explicitly noticed (even when there is no actual correlation; Chapman & Chapman, 1967) because they are consistent with people's existing background knowledge. The latter possibility appears more likely because of the computational complexity problem noted earlier. With the exception of extremely high or perceptually salient correlations (e.g., all tires being black), it would be difficult to notice correlations without any guidance of one's background knowledge. The present experiments are not informative in relation to this issue, because we examined natural concepts that people already possessed and did not examine the processes by which they had come to be aware of certain feature pairs. A future study using artificial stimuli with structures similar to natural categories will further shed light on this issue.

REFERENCES

Ahn, W., & Dennis, M. (2001). Dissociation between categorization and similarity judgment. In M. Ramscar & U. Hahn (Eds.), *Similarity and categorization* (pp. 87-107). New York: Cambridge University Press.

Ahn, W., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology*, **41**, 1-55.

Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, **91**, 112-149.

Barrett, S. E., Abdi, H., Murphy, G. L., & Gallagher, J. M. (1993). Theory-based correlations and their role in children's concepts. *Child Development*, **64**, 1595-1616.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: Plenum.

Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, **72**, 193-204.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instruments, & Computers*, **25**, 257-271.

Gluck, M. A. (1991). Stimulus generalization and representation in adaptive network models of category learning. *Psychological Science*, **2**, 50-55.

Gluck, M. A., & Bower, G. H. (1988). Evaluating an adaptive network model of human learning. *Journal of Memory & Language*, **27**, 166-195.

Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, **88**, 197-227.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Lassaline, M. E. (1996). Structural alignment in induction and similarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 754-770.

Malt, B. C., & Smith, E. E. (1984). Correlated properties in natural categories. *Journal of Verbal Learning & Verbal Behavior*, **23**, 250-269.

Medin, D. L. (1983). Structural principles in categorization. In T. J. Tighe & B. E. Shepp (Eds.), *Perception, cognition, and development: Interactional analyses* (pp. 203-230). Hillsdale, NJ: Erlbaum.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 37-50.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-196). New York: Cambridge University Press.

Murphy, G. L. (2000). Explanatory concepts. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 361-392). Cambridge, MA: MIT Press.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, **92**, 289-316.

Murphy, G. L., & Wisniewski, E. J. (1989). Feature correlations in conceptual representations. In G. Tiberghien (Ed.), *Advances in cognitive science: Vol. 2. Theory and applications* (pp. 23-45). Chichester, U.K.: Ellis Horwood.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.

Sloman, S., Love, B., & Ahn, W. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, **22**, 189-228.

Thagard, P. (1989). Explanatory coherence. *Behavioral & Brain Sciences*, **12**, 435-502.

Wattenmaker, W. D. (1991). Learning modes, feature correlations, and memory-based categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 908-923.

Wattenmaker, W. D. (1993). Incidental concept learning, feature frequency, and correlated properties. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 203-222.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

NOTES

1. Each property was weighted by the number of category members for which it appeared, and the family resemblance score for each item consisted of the sum of the weights of its properties.

2. It is also possible that sometimes people might first notice salient correlations and then subsequently develop theories about them. See the General Discussion section for further discussion.

3. We treat explanations and domain theories as the same in naive conceptual representations (Murphy, 2000) and use the terms interchangeably.

4. The term *implicit* has traditionally been used to designate abilities or knowledge that cannot be articulated yet could be demonstrated in an appropriate task. We used the same term here because participants failed to explicitly notice correlations inherent in their own representations when measured in a task employed in Part 1 of Experiment 1 in Malt and Smith (1984).

5. The exceptions to these will be discussed at the end of Experiment 2.

6. The participants were given differing numbers of explicit and implicit pairs, as defined by our Experiment 1, depending on which experimental form they were given. Fifteen participants received 90 explicit statements and 142 implicit statements, and 13 participants received 86 explicit statements and 144 implicit statements. These numbers might not appear to sum up to the total number of statements, but this is because some of the statements were presented to both groups of participants. The following breakdown gives the number of pairs in each category, with the parentheses representing the second form of the experiment: Bird—18 (17) explicit, 19 (19) implicit; Clothing,—4 (3) explicit, 17 (18) implicit; Flower—3 (4) explicit, 2 (5) implicit; Fruit—2 (2) explicit, 4 (3) implicit; Furniture—13 (13) explicit, 12 (12) implicit; Tree—5 (6) explicit, 15 (15) implicit.

7. This analysis does not capture cyclic dependency relations (e.g., A causes B and B also causes A) as theory-based relations because they will result in small difference scores. Yet, a maximum score should have captured this as a theoretical relation because, in this case, A depends on B, as well as B depending on A.

8. Negative relations were not used because, by definition, they will detrimentally lower typicality judgments. Malt and Smith (1984) do not describe exactly how negative correlations were treated in their Experiment 2. The most sensible thing to do is to negate one of the features to make the correlation positive, but that could make a pair identical to one of the other existing pairs (i.e., *can't fly*, *sits in trees* becomes *flies*, *sits in trees*, an already existing pair).

9. These weights were previously not published and were provided to the present authors by Barbara Malt.

APPENDIX A

The following table provides a list of all feature pairs rated as explicitly related in Experiment 1. The labels most often used for each feature pair in Experiment 3 are provided, along with the number of participants who used that label in parentheses.

Category	Feature 1	Feature 2	Mean Ratings on Relatedness	Commonly Used Labels
Bird	has feathers	is a bird	2.90	precondition (10), causes/allows (4)
	has a beak	is a bird	2.75	causes/allows (9), precondition (8)
	is a bird	lays eggs	2.65	causes/allows (12)
	is large	has large wings*	2.35	causes/allows (10)
	is large	has a large beak	2.20	causes/allows (8)
	flies	sits in trees*	2.10	causes/allows (15)
	has a large beak	eats fish*	1.95	causes/allows (17)
	eats fish	lives near the ocean*	1.90	causes/allows (15)
	has feathers	has a beak	1.90	causes/allows (6)
	has legs	sits in trees	1.75	causes/allows (8), precondition (4)
	has large wings	has a large beak	1.70	increases (5)
	has a beak	lays eggs	1.70	causes/allows (5)
	has feathers	lays eggs	1.65	*no common label*
	is small	eats insects	1.60	causes/allows (6)
	is small	has a large beak	−1.60	decreases/minimizes (7)
	is small	is large	−1.75	*no common label*
	is small	has large wings	−1.80	decreases/minimizes (9)
	can't fly	flies	−2.05	disables (6)
	can't fly	sits in trees	−2.40	decreases/minimizes (8), disables (4)
Clothing	is wool	is warm*	2.50	causes/allows (7), increases (6), subset/example (5)
	is clothing	is material*	2.20	subset/example (7), precondition (4)
	is warm	is worn in bad weather	2.15	causes/allows (10)
	is cotton	is comfortable*	2.00	causes/allows (10), increases (4)
	is cotton	is material*	1.65	subset/example (16)
Flower	has a long stem	is tall*	2.40	causes/allows (17)
	is small	has a long stem	−1.80	decreases/minimizes (6)
	is tall	is small	−1.80	*no common label*
Fruit	is sweet	tastes good	2.35	causes/allows (15)
	is juicy	tastes good*	1.90	causes/allows (11), increases (6)
Furniture	is soft	is comfortable*	2.60	causes/allows (7)
	has cushions	is for sitting on*	2.40	causes/allows (11), subset/example (4)
	has cushions	is soft*	2.40	causes/allows (11), increases (4)
	has cushions	is comfortable*	2.15	causes/allows (14)
	is comfortable	is for sitting on*	2.15	causes/allows (9), increases (4)
	is soft	is for sitting on*	2.00	causes/allows (6), increases (4), precondition (4)
	has springs	has cushions	1.90	causes/allows (8), precondition (4)
	has a cord	is electric*	1.70	causes/allows (11)
	has springs	is comfortable*	1.60	causes/allows (9), increases (4)
	has springs	is for sitting on*	1.55	causes/allows (8)
	has cushions	is on the wall	−1.50	*no common label*
	is comfortable	is on the wall	−1.60	*no common label*
	is for sitting on	is on the wall	−1.70	*no common label*
	is on the wall	is with a chair	−1.75	*no common label*
	is metal	is soft	−2.70	decreases/minimizes (7)
Tree	is big	is shady*	2.30	causes/allows (11), increases (5)
	has branches	is shady*	2.20	causes/allows (15)
	has a trunk	has bark*	2.15	causes/allows (8), precondition (5)
	is tall	is big*	2.15	causes/allows (8), subset/example (7)
	is old	is tall*	1.90	causes/allows (10)
	is old	is big*	1.85	causes/allows (11), increases (4)
	has needles	is green	1.65	causes/allows (9)

Note—Labels were condensed into similar groups for ease in interpretation. These groups were collapsed as follows: precondition, necessary, needs, requires as *precondition*; example, subset, type as *example/subset*; decreases, minimizes as *decreases/minimizes*; disables, inhibits, prevents as *disables*; causes, enables, allows, because, is due to, depends on, in order to, result of, determines, affects as *causes/allows*. The label *increases* did not include multiple labels. *These pairs were used for the construction of exemplars in Experiment 4.

APPENDIX B
Constant Features of Six Categories Used in Experiment 4

Category	Constant Features
Bird	has feathers, has claws, lays eggs
Clothing	can be bought, can be made, is used by humans
Tree	has leaves, manufactures chlorophyll, grows outdoors
Furniture	is found in houses, is man-made, can be bought in department stores
Fruit	is nutritious, is grown in warm climates, is eaten
Flower	has petals, grows in soil, is alive

APPENDIX C
Implicit Pairs Used in Experiment 4

Bird
- has a beak–is small
- has feathers–is black
- is gray/white–eats fish
- has a large beak–has legs

Clothing
- is worn in bad weather–is comfortable
- has a zipper–is colored
- has buttons–is colored
- has long sleeves–is colored

Fruit
- tastes good–is a summer fruit

Flower
- is yellow–has many petals

Tree
- grows in a warm climate–has bark
- grows in a warm climate–has a trunk
- is shady–has a trunk
- has bark–is tall
- has bark–is big
- has a trunk–is old

Furniture
- is metal–has a dial
- has a flat top–is with a chair
- has four legs–is comfortable
- is metal–is electric
- has a dial–is electric
- is wood–has a flat top
- has a dial–has a cord
- has four legs–is for writing/working
- has four legs–is with a chair

(Manuscript received October 23, 2000;
revision accepted for publication October 9, 2001.)