Abstract Coherent Categories

Bob Rehder
New York University

Brian H. Ross
University of Illinois at Urbana–Champaign

Many studies have demonstrated the importance of the knowledge that interrelates features in people’s mental representation of categories and that makes our conception of categories coherent. This article focuses on abstract coherent categories, coherent categories that are also abstract because they are defined by relations independently of any features. Four experiments demonstrate that abstract coherent categories are learned more easily than control categories with identical features and statistical structure, and also that participants induced an abstract representation of the category by granting category membership to exemplars with completely novel features. The authors argue that the human conceptual system is heavily populated with abstract coherent concepts, including conceptions of social groups, societal institutions, legal, political, and military scenarios, and many superordinate categories, such as classes of natural kinds.

Creating new categories of objects in the world depends on the theoretical, causal, and explanatory knowledge we possess about how the world works, not just how it appears. In this article we focus on a variety of such knowledge-based categories we refer to as abstract coherent categories. As we describe in detail below, such categories are coherent because they make sense in light of prior knowledge. They are abstract because they are independent of any fixed set of features or stimulus dimensions.

The important role that prior knowledge plays in human conceptual representations of categories has now been established in many category-related tasks, including classification, conceptual combination, and learning, to name a few (Keil, 1989; Malt & Smith, 1984; Medin & Shoben, 1988; Murphy, 1988; Murphy & Allopenna, 1994; Nakamura, 1985; Rips, 1989; Wattenmaker, Dewey, Murphy, & Medin, 1986; Wisniewski, 1995; Wisniewski & Medin, 1994; see Heit, 1997; Murphy, 1993, for reviews). From these findings, a broad consensus has emerged that the role of knowledge is to interrelate or link the concrete properties or features of a category to each other (Ahn, 1998; Keil, 1989; Rehder & Hastie, 2001). The role of knowledge in interrelating features is manifested explicitly in computational models that purport to represent prior knowledge (Pazzani, 1991; Rehder, 1999; Sloman, Love, & Ahn, 1998), including connectionist models (e.g., Choi, McDaniel, & Busemeyer, 1993; Heit & Bott, 2000; Rehder & Murphy, in press). Categories imbedded with theoretical or causal knowledge have been referred to as coherent because their attributes go together in light of that knowledge (Murphy & Medin, 1985).

According to this now standard account, representations of categories can be decomposed into their features and also their relations among features, and forming a representation of a new category consists of recruiting theoretical or causal knowledge that one already possesses to interrelate the new category’s features. Indeed, the presence of interattributive relations in people’s conceptions of natural categories has been amply demonstrated by empirical research. For example, Malt and Smith (1984) found that novel objects were rated as more typical of natural categories when they possessed pairs of features known by the participants to be related or correlated (e.g., flying and sitting in trees for birds). More recently, Ahn (1998, Experiments 1 and 2) found that undergraduates readily produced similar ratings of the strength of causal connections between attributes of natural categories ("goats give milk because they have four legs," also see Sloman et al., 1998). Thus, there appears to be little doubt that people’s conceptual representations include the presence of knowledge-linking category attributes.

The view that knowledge already possessed by a person provides the raw materials out of which the representation of a new coherent category is constructed (Gelman & Kalish, 1993) has also received support from experimental research. Pazzani (1991) found that the learning of a disjunctive category was facilitated when the to-be-predicted category was related by prior causal knowledge to the disjunction (i.e., a balloon is being stretched or is being held by an adult predicts that the balloon gets inflated) as compared with when no causal knowledge is present. Wattenmaker et al. (1986, Experiment 1) showed that participants learned to discriminate members of two categories more readily when category features could be interrelated by a personality trait (e.g., honest) compared with when no such underlying theme was present, even when the categories were given meaningless labels (A and B). Finally, Murphy and Allopenna (1994) also found that

Bob Rehder, Department of Psychology, New York University; Brian H. Ross, Beckman Institute of Advanced Science and Technology, University of Illinois at Urbana–Champaign.

This research was supported by National Science Foundation Grant SBR 97–20304. We thank Dorrit Billman, Gregory L. Murphy, Cynthia M. Sifonis, Edward J. Wisniewski, and two anonymous reviewers for their comments on earlier versions of this article and Angela Castellani for her assistance in conducting experimental sessions.

Correspondence concerning this article should be addressed to Bob Rehder, Department of Psychology, New York University, 6 Washington Place, New York, New York 10003. Electronic mail may be sent to bob.rehder@nyu.edu.
two categories with meaningless labels were learned much faster when their features could be united by an underlying theme (e.g., features such as “drives on glaciers,” “made in Norway,” and “heavily insulated” are united by the theme “arctic vehicle”) as compared with a control group whose category features were not related by prior knowledge. This final study provides especially convincing evidence of the importance of the relations among features supplied by prior knowledge because, whereas participants in the Pazzani and Wattenmaker et al. (1986) studies may not have been learning entirely new concepts (because they already knew the concept of honesty, and that stretching balloons predicts balloon inflation), Murphy and Allopenna used categories that were rated as novel by an independent group of participants.

Although an emphasis on the role of knowledge in interrelating category features represents a significant advance in our understanding of many real-world categories, it is nonetheless tied to a view that representations of categories include a specific set of concrete features, features that must appear with each category exemplar either necessarily (e.g., defining features of classically defined categories) or probabilistically (e.g., characteristic features of family-resemblance categories). In fact, however, many important types of conceptual structures do not specify any concrete features at all. Rather, some categories are defined by systems of relations that interconnect the features of category members without specifying what those features may be. Membership in such categories is determined by the extent to which a particular combination of features satisfies the system of relations. To emphasize both the absence of concrete features and the meaningfulness afforded by the presence of knowledge, we refer to such mental entities as abstract coherent categories. As we argue below, abstract coherent categories make up a significant portion of our conceptual world, including classes of natural kinds (e.g., mammals), social groups (families), political and military scenarios (revolutions, invasions), legal concepts (robbery, divorce), and types of societal institutions (governments), to name a few. We argue that many superordinate categories are abstract coherent categories.

In the left-hand side of Table 1, we present three exemplars of the abstract coherent category, called morkels, used in the present research. One should note that each of the three examples describes a coherent device. For example, a device that gathers spilled oil on the surface of water with a sponge is coherent, because spilled oil is often the result of water-borne shipping accidents, because oil floats on water, and because sponges are good for soaking up fluids (such as oil). However, the coherent morkels of Table 1 are not merely a collection of unrelated coherent devices. Rather, the coherent morkels also exhibit coherence as a group because they share a common set of abstract interattributive relationships. For example, an apt description of the coherent morkels would be “machines that gather a certain type of pollution, in an area where such pollution is likely to be found, with a device suitable for collection of that type of pollution.” The morkel category exhibits (abstract) coherence and does so without its members sharing any concrete features. Moreover, not only do the three examples of morkels in Table 1 have no features in common, the coherent morkel category appears to admit of a virtually infinite number of machines that can be construed as pollution-cleaning devices, none of which need share any concrete features with any other. For example, our intuitions suggest that after learning that the three coherent machines of Table 1 are called morkels, category learners would be likely to think that a new machine that “operates in highway tunnels,” “works to remove carbon monoxide,” and “has large intake fans” is a morkel as well, even though it shares none of the concrete features of the initial three examples of coherent morkels.

In contrast, the right-hand side of Table 1 lists three exemplars of an incoherent morkels category, formed from the same features of the coherent morkels, but with the features recombined to form new machines. This recombining of features has reduced this second morkel category to a meaningless collection of senseless objects, such as “floats in the stratosphere,” “works on absorbing spilled oil,” and “has a shovel.” In contrast to the coherent morkel category in which explanatory reasoning might lead one to conclude that “operates in highway tunnels,” “works to remove carbon monoxide,” and “has large intake fans” is also a morkel, such reasoning would appear to be of no help in deciding whether this new machine was a member of the incoherent morkel category, because incoherent morkels are not bound together by abstract coherence relationships. Assuming that simple similarity becomes the default criterion for judging category membership in the absence of explanatory knowledge, we suspect most categorizers would be unlikely to include the new machine in the incoherent morkel category, because it possesses none of the features of the three known category members.

There has been little study of abstract coherent concepts. Wattenmaker et al. (1986, Experiment 3) taught participants to identify members of a “house painter” category that can be construed as abstract because category members belonged to one of

<table>
<thead>
<tr>
<th>Coherent morkels</th>
<th>Incoherent morkels</th>
</tr>
</thead>
<tbody>
<tr>
<td>operates on surface of water</td>
<td>floats in the stratosphere</td>
</tr>
<tr>
<td>works on absorbing spilled oil</td>
<td>works on absorbing spilled oil</td>
</tr>
<tr>
<td>coated with spongy material</td>
<td>has a shovel</td>
</tr>
<tr>
<td>operates on land</td>
<td>operates on surface of water</td>
</tr>
<tr>
<td>works to gather harmful solids</td>
<td>works to gather harmful solids</td>
</tr>
<tr>
<td>has a shovel</td>
<td>has an electrostatic filter</td>
</tr>
<tr>
<td>floats in the stratosphere</td>
<td>operates on land</td>
</tr>
<tr>
<td>works to absorb dangerous gaseous ions</td>
<td>works to absorb dangerous gaseous ions</td>
</tr>
<tr>
<td>has an electrostatic filter</td>
<td>coated with spongy material</td>
</tr>
</tbody>
</table>
two house-painter subtypes (indoor painters and outdoor painters) that shared no concrete features ("works inside" and "works year round" vs. "works outside" and "doesn't work in winter"). Similarly, Wisniewski and Medin (1994) found that participants used abstract concepts such as detailed, as indicated by their use of a wide variety of concrete features, as evidence of detail in children's drawings (e.g., buttons, shirt collars). (In that study, abstract concepts such as detailed were used by participants for evidence of more complex categories; e.g., "drawings by creative children"). However, in each of these studies, the abstract concept (house painter or detailed) was likely to have already been known to participants. An experiment that may have involved the learning of a novel abstract category was reported by Wisniewski (1995, Experiment 4). Participants were presented with members of two artifact categories, described as "captures animals" (and also given the label platypus) and "cleans up pollution" (mornes). (Our own pollution-cleaning materials were adapted from those used by Wisniewski, 1995.) For example, members of the platypus category included an instance that "contains acorns," "caught a squirrel," and "has a small metal container," and also an instance that "contains peanuts," "caught an elephant," and "has a very large metal container." Wisniewski intended the platypus category to be an abstract coherent category, because it is defined by a coherent set of relations but no concrete features. Specifically, platypus are devices for capturing animals of Type X using bait that X's like to eat and possessing containers suitable for holding X's once caught. On a transfer test, Wisniewski found that novel combinations of features are more coherent ("contains peanuts" and "caught a squirrel") were considered better platypus than noncoherent combinations ("contains acorns" and "caught an elephant"). Because the constituent features of each novel combination were equated for their functional relevance to and statistical diagnosticity of the platypus category, Wisniewski attributed this result to the presence of interfeature relationships supplied by participants' prior knowledge concerning the eating preferences of squirrels and elephants.

Although important, we believe that Wisniewski's (1995) demonstration fails to capture the typical learning situation faced by people acquiring new abstract categories because he provided participants with category descriptions ("captures animals" and "pollution-cleaning devices") that specified the to-be-learned abstractions. However, the remarkable ability of language to induce novel and abstract mental concepts (as demonstrated, e.g., by Barsalou's, 1983, 1985, ad hoc categories; e.g., "things to take out of one's house in case of fire") is not in doubt. The more common learning problem faced by children (and often adults) is that a new concept's label is opaque, in that it provides no information about the category. A person's task is usually to induce new abstractions rather than to merely apply ones that are directly provided. Thus, the acquisition of an abstract coherent category remains underemphasized in experimental research.

The goal of the present study was to examine the conditions under which people acquire abstract coherent concepts that are truly novel, and to do so in the absence of linguistic descriptions that, in effect, presuppose the abstraction to be learned. To this end, across a series of experiments, participants learned about either the coherent or incoherent morkels shown in Table 1. Like Pazzani (1991), Wattenmaker et al. (1986), and Murphy and Alloppenna (1994), we will take facilitated learning in the coherent versus the incoherent condition as evidence for the importance of the inter-

attribute relations, because exactly the same features are used in the two conditions.

Across experiments, we varied the extent to which the learning task either did or did not focus category learners on relationships among category attributes, so as to discover those abstract coherence relationships. We began our investigation by using a learning task that we felt would be most conducive to learning abstract coherent categories. In Experiment 1, participants learned about either coherent or incoherent morkels by predicting missing features. We expected that this procedure would focus participants on the relationships among category attributes (Yamalrich & Markman, 1998) and would lead participants in the coherent condition to induce the abstract coherence relations that we intended them to. In Experiment 2, we used a classification-learning task in which participants were required to distinguish morkels (either coherent or incoherent) from an incoherent contrast category (nonmorkels). We expected that the classification task would be a more stringent test of participants' ability to induce the abstract morkel category, because it would draw attention away from within-category attribute relationships and toward between-category differences. In Experiment 3, we further explored the dependence of abstract category induction on the learning task by noting that the categories in Experiment 2 were labeled asymmetrically (morkels vs. nonmorkels) and that this asymmetry might assist learners in discovering the coherent morkels' meaningful interattribute relationships (Goldstone, 1996). To test this possibility, we also used a classification-learning task in Experiment 3, but the incoherent contrast category was given its own label (krenshaws).

To foreshadow the results of the first three experiments, participants in the coherent condition exhibited accelerated learning and superior transfer performance compared with those in the incoherent condition. These results were obtained even though morkels were a novel category and even though no linguistic description of the category was provided. However, by themselves, these findings do not demonstrate that participants in fact induced the abstract coherent category that we intended them to, because coherent participants might have simply learned to associate three concrete (albeit coherent) subcategories (water, air, land morkels) with one category label (morkel). To test this alternative hypothesis, we once again used a classification-learning task in Experiment 4, but we also directly tested the level of abstraction of participants' mental representation of morkels by asking them to classify a series of novel machines (with novel features), some of which were coherent pollution-cleaning devices.

Experiment 1

In Experiment 1, we had participants focus on relations among category attributes by having them perform a feature-prediction task. Category exemplars with one missing feature were presented, and participants were required to predict that feature. In the coherent condition, category exemplars consisted of the coherent pollution-cleaning morkels shown in Table 1. We increased the number of exemplars of this category to nine by adding values on two additional dimensions to the three exemplars of Table 1. That is, the coherent morkels consisted of three instances of the water, land, and air morkels in Table 1, but each instance was made unique by also describing the material it was made of and its power source. For example, a morkel that gathers spilled oil on water
with a sponge was described sometimes as being made of aluminum and powered by battery, sometimes as being made of steel and powered by methane, and sometimes as being made of copper and powered by alcohol. The values for these additional dimensions were deliberately chosen to not cohere with values of the first three dimensions. The formal category structure used in Experiment 1 for coherent morkels is presented in Table 2; the concrete dimension values used to instantiate Table 2's abstract structure are presented in Table 3.

Similarly, in the incoherent condition, the incoherent morkels shown in Table 1 were used, with the same two additional dimensions (composition and power source) added to construct a total of nine exemplars from the three incoherent morkels of Table 1. The formal structure of this incoherent morkel category is also shown in Table 2.

On formal grounds, the structures of the coherent and incoherent morkels shown in Table 2 are identical. For both categories, all five dimensions have three values, each appearing with a one-third probability. Furthermore, Dimensions A, B, and C are perfectly correlated. For example, when Dimension A takes on Value 1, Dimensions B and C always take on Value 1 for the coherent morkels and Values 2 and 3, respectively, for the incoherent morkels. Finally, Dimensions A, B, and C are perfectly uncorrelated with the additional Dimensions D and E, and Dimensions D and E are perfectly uncorrelated with one another.

In this experiment’s inference-learning task, the missing-to-be-predicted feature always occurred on either Dimension A, B, or C. The correlational structure of Dimensions A, B, and C for both the coherent and incoherent categories means that participants could correctly predict a missing feature on one of these dimensions from either of the other two. Nevertheless, we predicted that participants in the coherent group would learn to correctly predict missing features more readily than those in the incoherence group, because the coherence relations between Dimensions A, B, and C of coherent morkels would promote the learning of these dimensions in relation to the other two.

Success on this inference-learning task does not require complete knowledge of the co-occurrences of values on Dimensions A, B, and C. For example, perfect feature-prediction performance can be achieved by learning which attribute values co-occur on Dimensions A and B and A and C but not on Dimensions B and C. However, we predicted that participants in the coherent condition would not just learn to predict missing features more readily but also would apprehend the three coherent morkel subtypes; that is, the water, land, and air morkels shown in the left-hand side of Table 1. To determine whether participants learned of all the co-occurrences of values among the three correlated dimensions (i.e., whether they learned the three morkel subtypes) after inference learning, we asked participants to perform a classification task in which they judged the category membership of exemplars possessing only two values from Dimensions A, B, or C. Suboptimal classification performance on such exemplars will be indicative of a failure to learn the three subtypes of morkels.

**Table 2**  
Abstract Structure of the Coherent and Incoherent Morkel Categories

<table>
<thead>
<tr>
<th>Exemplar</th>
<th>Dimension</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coherent morkels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1*</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1*</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1*</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2*</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>2*</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>2</td>
<td>2*</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3*</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>3*</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>3</td>
<td>3*</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Incoherent morkels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1*</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2*</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2*</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3*</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>1*</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>1</td>
<td>2*</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2*</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>3*</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>3</td>
<td>3*</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

*Features listed as missing in Experiment 1.

**Method**

**Participants.** Twenty-four University of Illinois at Urbana-Champaign undergraduates received either course credit or $5 for participating in this experiment.

**Materials.** During the inference-learning phase of Experiment 1, participants were presented with either the coherent or incoherent morkels from Table 2. Each morkel was described as possessing four features; the fifth missing feature was designated with "??" and was the feature that participants were required to predict. Table 2 indicates which features were missing from each morkel during inference learning. These morkels were presented to participants on individual 3 × 5-in. index cards. Dimension values were listed on the index cards in a consistent order: Dimension A, B, C, D, then E.

During the subsequent classification test, participants were asked to categorize 18 two-feature exemplars, 9 from the category observed during inference learning, and 9 from a previously unexperienced contrast category. The abstract structure of the contrast category is presented in Table 4. This contrast category consists of the same features as both the coherent and incoherent morkels but recombined in yet a third way to yield exemplars distinct from those of either morkel category. In the coherent condition, participants were asked to classify the nine examples of coherent morkels that can be constructed with two features on Dimensions A, B, or C (i.e., 11xx, 22xx, 33xx, 1x1xx, 2x2xx, 3x3xx, x11xx, x22xx, x33xx; x denotes a missing value) and the corresponding nine two-feature exemplars from the contrast category (i.e., 13xx, 21xx, 32xx, 1x2xx, 2x3xx, 3x1xx, x32xx, x13xx, x21xx). In the incoherent condition, participants were asked to classify the nine two-feature exemplars of incoherent morkels (i.e., 12xx, 31xx, 23xx, 13xx, 32xx, 21xx, 23xx, 1x2xx, x12xx, x31xx) and the same nine two-feature exemplars from the contrast category used in the coherent condition. (Values assigned to Dimensions D and E of exemplars of the contrast category in Table 4 were used in subsequent experiments.) These two-feature exemplars were presented to participants on individual 3 × 5-in. index cards. Missing features were not marked with a "??" but rather each card just listed two values. Dimension values were listed on the index cards in the order of Dimension A, B, then C.

**Procedure.** The experimental session commenced with participants being told that they would be learning about types of machines called
**Table 3**

Values Assigned to Dimensions of the Abstract Category Structures Presented in Tables 2 and 4

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>operates on surface of water</td>
<td>operates on land</td>
<td>floats in the stratosphere</td>
</tr>
<tr>
<td>B</td>
<td>works on absorbing spilled oil</td>
<td>works to gather harmful solids</td>
<td>works to absorb dangerous gaseous ions</td>
</tr>
<tr>
<td>C</td>
<td>coated with spongy material</td>
<td>has a shovel</td>
<td>has an electrostatic filter</td>
</tr>
<tr>
<td>D</td>
<td>made of aluminum</td>
<td>made of steel</td>
<td>made of copper</td>
</tr>
<tr>
<td>E</td>
<td>powered by battery</td>
<td>powered by methane</td>
<td>powered by alcohol</td>
</tr>
</tbody>
</table>

**Results**

*Inference learning.* Overall, the coherent group learned to predict features of morkels much more readily than the incoherent group. For example, all 12 coherent participants reached the criterion of 0 errors per block, doing so in an average 1.75 blocks. In contrast, only 9 of the 12 incoherent participants reached criterion in the maximum of 8 blocks. Under the (most conservative) assumption that the 3 incoherent participants who did not reach criterion would have done so on the very next (i.e., ninth) block, the average number of blocks to criterion for the incoherent group was 5.75. The difference between the two groups on the number of blocks to criterion was highly significant, $t(22) = 5.06, p < .0001$.

The relative ease with which the coherent group learned to predict missing features of morkels was also reflected in the total number of errors committed. Whereas participants in the incoherent group committed an average of 17.5 errors during inference learning, the average coherent participant committed only 1.0 error, $t(22) = 5.65, p < .0001$. Of interest, 3 of the 12 coherent participants made no errors at all. Apparently, error-correcting feedback was not required for coherent participants to notice the coherence between the features of a presented morkel and the correct choice for its missing feature.

*Two-feature classification test.* The purpose of the two-feature classification test was to determine how thoroughly participants learned the co-occurrences of values on Dimensions A, B, and C. In fact, 11 of the 12 coherent participants performed perfectly on this test, classifying all 18 of the two-feature exemplars correctly. The remaining coherent participant committed three errors. In contrast, only two members of the incoherent group performed perfectly, and the average number of errors committed by participants in this group was 3.8. The difference in error rates between groups was highly significant, $t(22) = 4.21, p < .0001$.

One possible reason for these differences is that the two groups are not equated on the degree of learning. Whereas all 12 of the coherent participants reached the learning criterion, only 9 of the 12 incoherent participants did so. Group differences on the two-feature classification test may conceivably be due to the 3 nonlearning incoherent participants. However, even when those 3 participants are excluded from the analysis, the incoherent group still performed more poorly than the coherent group, committing an average of 3.4 errors, $t(19) = 3.63, p < .01$.

An additional analysis using the confidence ratings was carried out. An adjusted confidence score was calculated by negating the
confidence ratings on those trials in which the classification decision was incorrect, and then summing the ratings over the 18 trials. Maximum performance on the classification test (i.e., correctly classifying all 18 exemplars with the maximum confidence rating of 7) results in an adjusted confidence score of 126; chance responding would yield a score of 0. The average adjusted confidence scores were 115.5 and 73.3, respectively, for the 12 coherent and 9 incoherent participants who reached the learning criterion. Thus, consistent with the number of errors, the adjusted confidence score reveals that the coherent participants performed almost perfectly on the two-feature classification test and much better than the incoherent participants, \( t(19) = 3.23, p < .01 \).

Although the 9 incoherent participants who reached criterion performed worse than those in the coherent group on the two-feature classification test, they demonstrated that they were able to use their feature-prediction knowledge to correctly classify some transfer items, as both their average error rates (3.4) and adjusted confidence scores (73.3) were reliably different than chance (9 and 0, respectively), \( t(8) = 5.71, p < .001 \), and \( t(8) = 5.54, p < .001 \).

**Discussion**

It would be difficult to overstate the ease with which participants learned to predict missing features of exemplars of the coherent morkel category in Experiment 1. On the one hand, it might seem that the inference-learning task was not especially difficult, because each dimension on which missing features appeared (A, B, or C) was perfectly correlated with two of these other dimensions. On the other hand, it was still necessary for participants to learn which values on Dimensions A, B, and C co-occurred, and that Dimensions D and E were not useful in predicting the missing features. Nevertheless, participants in the coherent condition achieved perfect feature-prediction performance, receiving feedback on an average of only one trial, and 11 of 12 participants then went on to perform perfectly on the two-feature classification test. Apparently, coherent participants achieved near-flawless performances by spontaneously apprehending the meaningful relationships obtained among attributes of the exemplars of the coherent morkel category.

Despite the absence of meaningful relationships between category attributes, participants in the incoherent condition also experienced considerable success learning to predict missing features of the incoherent morkel category and then used this knowledge to correctly classify the two-feature transfer items. Nevertheless, the learning and classification performance of the incoherent group was dramatically worse than the coherent group. Differences between the groups on the two-feature classification test were obtained even when the data from the three nonlearning incoherent participants were excluded, indicating that the information contained in the mental representation of the morkel category differed between the coherent and incoherent conditions even when both groups were equally good (in fact, perfect) at predicting missing features. That is, the presence of interfature relationships in the coherent condition not only affected how fast learning proceeded but also how much about the statistical structure of the category was learned.

Although we are impressed by the learning and classification performance achieved by the coherent group, this performance was achieved in an inference-learning task that was deliberately chosen to maximize the probability that those participants would notice and use the meaningful interattribute relations. Further, missing features always appeared on exactly those dimensions that were mutually coherent (Dimensions A, B, or C), making it even more likely that learners would apprehend those relations. The possibility exists that the advantage exhibited by the coherent group might be attenuated (or eliminated) under different learning conditions. To test this possibility, we investigated the effect of abstract coherence using a traditional category learning paradigm in Experiment 2.

**Experiment 2**

In Experiment 2, participants learned to distinguish morkels (coherent or incoherent) from exemplars of the contrast category shown in Table 4. This contrast category consists of the same features as both the coherent and incoherent morkels but recombined in yet a third way to yield exemplars distinct from those of either morkel category. Because the combinations of features specified by this contrast category exhibit no interattribute coherence relations, the task for participants in the coherent group was to learn to distinguish one coherent category (morkels) from one incoherent contrast category (nonmorkels), whereas the task for the incoherent group was to learn to distinguish members of two equally incoherent categories.

Our expectation was that the classification-learning task would lead learners to initially focus more on between-category differences rather than on within-category attribute relationships. However, successfully distinguishing morkels from nonmorkels requires participants to learn something about the interattribute correlations, because individual attribute values were not diagnostic of category membership. For example, devices that "operate on the surface of the water" appear equally often as morkels (coherent or incoherent) and nonmorkels. We predicted that the interattribute coherence relations among attributes of coherent morkels would facilitate learning of this critical interattribute correlational information as compared with the incoherent morkels.

As was the case in Experiment 1, successful learning in Experiment 2 did not require participants to have complete knowledge of all the co-occurrences of features on Dimensions A, B, and C. Thus, after learning, we once again asked participants to classify a series of exemplars that possessed two features from those dimensions in order to determine the extent to which learners encoded all the co-occurrences of values, that is, the three morkel subtypes.

**Method**

**Participants.** Twenty-four University of Illinois at Urbana-Champaign undergraduates received course credit for participating in this experiment.

**Materials.** For the classification-learning task, 18 3 × 5-in. index cards listing all five attributes of an exemplar were prepared for both the coherent and incoherent groups. For the coherent group, these cards described the nine coherent morkel exemplars from Table 2, and the nine contrast category exemplars from Table 4. For the incoherent group, these cards described the nine incoherent morkel exemplars from Table 2, and the same nine contrast category exemplars. Unlike Experiment 1, these exemplars possessed no missing attributes, that is, all five features were present. For the subsequent classification test, we used the same 18 two-feature items used in Experiment 1.
Procedure. During classification learning, participants were presented with each of the 18 five-feature exemplars and were asked to classify each. Immediate feedback was provided after every categorization decision. Participants were told that they could study each card as long as they wished after receiving feedback. Blocks were repeated until the participant made no more than two errors or reached the maximum number of blocks of four. The order of presentation for morkels within each block was randomized. At the end of the experiment, participants once again described morkels. Each session lasted approximately 30 min.

Results

Classification learning. Nine of the 12 coherent participants reached the learning criterion of two or fewer errors per block in 4 blocks or fewer. In comparison, only 3 of the 12 incoherent participants reached criterion. As in Experiment 1, the conservative assumption that nolearners would reach criterion on the very next (fifth) block was made. Under this assumption, the mean number of blocks to criterion was 3.08 and 4.42 for the coherent and incoherent groups, respectively, t(22) = 2.72, p < .05. The faster learning exhibited by the coherent group was also reflected in the total number of errors. The coherent group committed a total of 18.5 errors, which was significantly less than 29.5 errors for the incoherent group, t(22) = 2.10, p < .05.

Two-feature classification test. On the classification test, the coherent group committed an average of 3.7 errors on the 18 two-feature items and achieved an average adjusted confidence score of 62.2. In contrast, the averages were 7.8 and 13.3, respectively, for the incoherent group; the groups were reliably different on each of these two measures: t(22) = 3.24, p < .01, for errors; t(22) = 3.44, p < .01, for confidence scores.

The poor performance of the incoherent group on the classification test can be attributed in part to the fact that only 3 of those 12 participants reached criterion on the classification-learning task. However, even when only the 9 coherent and 3 incoherent participants who reached the learning criterion are considered, significant differences between groups remain on both error rates (2.0 for the coherent group vs. 5.7 for the incoherent group), t(10) = 2.98, p < .05, and adjusted confidence scores (82.1 vs. 37.7), t(10) = 3.20, p < .05.

Discussion

The classification-learning task used in Experiment 2 required that participants learn some of the co-occurrences between values on Dimensions A, B, and C, because none of the individual values on those dimensions (or on Dimensions D and E) were diagnostic of category membership. Consistent with the results of Experiment 1, the current results show that the learning of these co-occurrences was facilitated by the presence of coherence relationships between category attributes in the coherent morkel category as compared with the incoherent morkel category.

The Murphy and Allopenna (1994) study discussed earlier is also an instance of background knowledge facilitating the learning of novel categories in a classification-learning task. However, in that study, the two contrasting categories were each coherent and concrete (e.g., jungle vehicles vs. arctic vehicles), whereas in the current study the morkel category was coherent and abstract (the water morkel, the land morkel, and the air morkel), and the contrast category exhibited no coherence. Further, the Murphy and Allopenna categories exhibited a family resemblance structure in which concrete category features were predictive of category membership, in comparison with the current experiment in which features possessed no such predictive value. Thus, the current experiment demonstrates facilitated learning brought about by background knowledge with very different category structures than those used by Murphy and Allopenna and other similar studies of knowledge and category learning.

For both the coherent and the incoherent groups, all that was required to discriminate between morkels and nonmorkels in Experiment 2 was to learn the co-occurrence of values on two of the three critical dimensions (A, B, or C), or to learn a series of isolated two-value fragments on all three dimensions. Indeed, the average of 5.7 errors committed by the incoherent group on the 18 two-feature classification test indicates that those participants did not learn the complete co-occurrence information on those dimensions. In contrast, the number of errors committed by the coherent learners (2.0) suggests that the coherence relations among category attributes led to participants learning most of the categories' statistical structures. This result is consistent with the participants learning about the three morkel subtypes in the coherent but not the incoherent condition.

Experiment 3

We attributed the slower learning performance in Experiment 2 versus Experiment 1 to the initial focus on between-category differences brought about by the classification-learning task as compared with an inference-learning task that naturally focuses attention on within-category attribute relations. Nevertheless, even in Experiment 2, participants may have had a tendency to focus on the internal structure of the morkel category because of our use of asymmetric category labels, morkels versus nonmorkels. Because real-world categories are often defined relative to other positively labeled categories (Goldstone, 1996), in Experiment 3 we asked whether the learning advantage exhibited by the coherent group would persist even when the two categories (morkels and the contrast category) received equal focus because the contrast category was given its own label.

Method

Participants. Twenty-four University of Illinois at Urbana-Champaign undergraduates received course credit for participating in this experiment.

Materials and procedure. The materials and procedure of Experiment 3 were identical to those of Experiment 2, except that the contrast category was given the label krenshaws.

Results

Classification learning. The classification-learning performance was much lower than in Experiment 2. Five of the 12 coherent participants reached the learning criterion of two errors per block in 4 blocks or less. In comparison, none of the 12 incoherent participants reached criterion. A Fisher's Exact Test revealed that significantly more coherent than incoherent participants reached the learning criterion (p < .001).

This learning exhibited by the coherent group was also reflected in the total number of errors. The coherent group committed a total
of 28.2 errors, which was significantly fewer than 35.5 for the incoherent group, \( t(22) = 2.16, p < .05 \).

**Two-feature classification test.** On the classification test, the coherent group committed an average of 6.6 errors on the 18 two-feature items and achieved an average adjusted confidence score of 29.2. This performance is markedly worse than in the first two experiments, presumably because, in this experiment, fewer coherent participants reached the learning criterion. When only the 5 coherent participants who reached criterion are considered, the average number of errors and adjusted confidence scores were 3.0 and 70.0, respectively. This performance was reliably better than chance (9 errors and 0 adjusted confidence score) on both measures, \( t(4) = 2.50, p < .05 \), and \( t(4) = 2.57, p < .05 \).

As expected, given that none of the incoherent participants reached the learning criterion, this group exhibited no learning on the two-feature classification test: error rate of 9.5 and adjusted confidence score of ~1.8, both measures slightly worse than chance.

**Discussion**

As expected, the equal focus brought about by the use of symmetrical category labels in Experiment 3 reduced the number of coherent participants who learned to distinguish markels from nonmarkels, from 9 (of 12) participants in Experiment 2 to 5 (of 12) participants in the current experiment. However, despite this slower learning in the current experiment, the coherent group continued to exhibit superior performance relative to the incoherent group. Thus, the presence of meaningful interattribute relationships facilitated the learning of the coherent morkel category in yet a third experimental test.

Those participants in the coherent condition that reached the learning criterion performed well on the two-feature classification test, suggesting that when they learned, they learned the entire set of value co-occurrences on Dimensions A, B, and C. As in Experiments 1 and 2, we suggest that this performance indicates that those participants learned the three morkel subtypes. We present additional evidence in favor of this interpretation in the following section.

The use of symmetrical category labels in Experiment 3 also hurt learning in the incoherent condition as compared with Experiment 2. It is surprising that the additional focus on the contrast category hurt learning, because that category possessed interattribute correlations that were as useful for discriminating the two categories as those possessed by the morkel category. Apparently, the spread of attention across two categories resulted in incoherent participants learning the correlational structure of neither, perhaps because of the excessive resource demands that such learning requires. In the General Discussion, we will consider more detailed explanations for the reduced learning experienced by both the coherent and incoherent groups as a result of symmetrical category labeling.

**Analysis of Debriefing Questions**

In Experiments 1, 2, and 3, the low error rates (and high adjusted confidence scores) achieved by the coherent participants on the two-feature classification tests indicated that those individuals had extracted most of the interattribute correlational information, even though not all of that information was necessary to reach criterion during inference learning (in Experiment 1) or classification learning (in Experiments 2 and 3). This result is consistent with the view that coherent participants apprehended the three concrete coherent morkel subtypes shown in the left side of Table 1. To obtain more direct evidence for this conclusion, we examined participants' responses to the debriefing question in which they were asked to describe markels. We took as evidence that a participant apprehended a subtype if he or she listed the three co-occurring features on Dimensions A, B, and C as a unit, with one point given for each subtype described in this manner. For example, when describing markels, one coherent participant whose response was given 3 points said “If on water, a morkel gathers oil and has a sponge. If on land, it gather solids and has a shovel. If in the air, it collects ions with a filter.” Similarly, the response of one incoherent participant whose response was given 3 points was “...if in water it cleans up solid waste and uses an electrostatic filter. On land, gaseous ions with a sponge. In stratosphere, picks up oil spills, uses a shovel.”

Considering only the 26 coherent learners and the 12 incoherent learners from Experiments 1, 2, and 3, the average number of subtype descriptions was 1.46 and 0.50, respectively, out of 3, \( t(35) = 2.11, p < .05 \). This result confirms that a major factor responsible for the excellent performance on the two-feature transfer test of the coherent groups was their apprehension of the three coherent morkel subtypes.

We also asked participants what markels were used for. We coded an answer to this question as abstract if they responded with “cleans up pollution” or “cleans up environment” (or synonymous phrases such as “removes dangerous stuff,” “disposes of waste,” “cleans up environment,” “cleans earth”). Somewhat surprisingly, the incoherent learners produced more abstract morkel descriptions (83%) than the coherent learners (65%), although this difference did not approach significance (\( z = 1.13, p > .20 \)). Also surprising is the fact that half of the nonlearners produced abstract morkel descriptions (30% and 58% of the coherent and incoherent participants, respectively); the learner–nonlearner difference was only marginally significant (71% vs. 50%, respectively; \( z = 1.83, p < .10 \)). Apparently, most participants concluded that markels had something to do with cleaning up pollution, perhaps by abstracting over features as “works on absorbing spilled oil,” “works to gather harmful solids,” and “works to absorb dangerous gaseous ions.” However, forming an abstraction over a single feature dimension is not equivalent to the abstract morkel category because the abstract interattribute coherence relationships are not specified; specifically, the meanings of gathering and the location must be appropriate to the type of pollution. The purpose of Experiment 4 was to directly test whether participants acquired the abstract coherent morkel category.

**Experiment 4**

The goal of Experiment 4 was to determine whether coherent participants learned the abstract coherent morkel category that we intended them to. Although the results of the two-feature classification test (combined with our analysis of the debriefing questions) indicate that coherent participants perceived the individual subtypes of water, air, and land markels, these participants may have simply represented markels as a disjunctive category in
which each morkel subtype was associated to the same category label, without noticing that such subtypes could be conceived of as members of one, abstract coherent category. Although our analysis of the debriefing questions indicated that most participants acquired some abstract morkel information, this information may have been related to just one feature dimension. To test these possibilities, Experiment 4 consisted of a replication of Experiment 2, but at the end of the experiment, participants were asked to rate the morkel category membership of a series of new machines that either could or could not be construed as coherent pollution-cleaning devices. Because these new machines were constructed from completely new features, simple associations between the three morkel subtypes and the category label morkels would provide no basis for deciding which of the new machines were morkels. Because the new incoherent machines were constructed from the same features as the new coherent machines, an abstraction over a single feature dimension would also be of no assistance in identifying the new morkels. An ability to identify new coherent pollution-cleaning devices as morkels will constitute evidence that participants acquired the abstract coherent morkel category.

**Method**

**Participants.** Thirty-two University of Illinois at Urbana-Champaign undergraduates received course credit for participating in this experiment.

This experiment used one-third more participants than the Experiments 1–3 because of the greater variability introduced by the use of novel transfer items.

**Materials.** We constructed six new machines that we believed described coherent pollution-cleaning devices. These novel machines are presented in the left-hand side of Table 5. The first three incoherent novel morkels on the right-hand side of Table 5 were constructed by recombining the values on Dimensions A, B and C of the first three coherent novel morkels. The last three incoherent novel morkels were constructed by recombining the values of the final three coherent novel morkels. Because the same features were used to construct the coherent and incoherent novel items, any differences in the morkel category membership ratings of those items cannot be attributed to differences in feature similarities between the two types of items.

To confirm that the six coherent novel morkels of Table 5 were indeed viewed by our participant population as coherent, and that the six incoherent novel morkels were viewed as incoherent, an independent group of 15 participants was asked to rate the plausibility of those machines on a scale from 1 (very implausible) to 7 (very plausible). Plausibility was defined to mean that “the machine’s features go together in some coherent, sensible way such that the machine is able to fulfill some useful function or purpose.” The results of this rating study confirmed that the coherent novel morkels of Table 5 were viewed as plausible, as the ratings for those six items were 6.0, 5.7, 6.5, 5.4, 6.1, and 6.5. The ratings for the six incoherent items were 1.3, 2.9, 3.0, 2.3, 2.9, and 4.0, indicating that those items were all considered clearly implausible, except for the sixth one, which received a plausibility rating in the middle of the 1–7 scale. Because this item was

<table>
<thead>
<tr>
<th>Coherent novel morkels</th>
<th>Incoherent novel morkels</th>
</tr>
</thead>
<tbody>
<tr>
<td>operates in highway tunnels</td>
<td>operates in highway tunnels</td>
</tr>
<tr>
<td>works to remove carbon monoxide</td>
<td>works to remove malaria-ridden mosquitoes</td>
</tr>
<tr>
<td>has large intake fans</td>
<td>has a large magnet</td>
</tr>
<tr>
<td>made of tin</td>
<td>made of tin</td>
</tr>
<tr>
<td>powered by diesel fuel</td>
<td>powered by diesel fuel</td>
</tr>
<tr>
<td>operates in swamps</td>
<td>operates in swamps</td>
</tr>
<tr>
<td>works to remove malaria-ridden mosquitoes</td>
<td>works to gather shards of metal</td>
</tr>
<tr>
<td>has a finely-woven net</td>
<td>has large intake fans</td>
</tr>
<tr>
<td>made of tungsten</td>
<td>made of tungsten</td>
</tr>
<tr>
<td>powered by butane</td>
<td>powered by butane</td>
</tr>
<tr>
<td>operates in war zones</td>
<td>operates in war zones</td>
</tr>
<tr>
<td>works to gather shards of metal</td>
<td>works to remove carbon monoxide</td>
</tr>
<tr>
<td>has a large magnet</td>
<td>has a finely-woven net</td>
</tr>
<tr>
<td>made of brass</td>
<td>made of brass</td>
</tr>
<tr>
<td>powered by kerosene</td>
<td>powered by kerosene</td>
</tr>
<tr>
<td>operates in parks</td>
<td>operates in parks</td>
</tr>
<tr>
<td>works to gather discarded paper</td>
<td>works to remove lost fishing nets</td>
</tr>
<tr>
<td>has a metal pole with a sharpened end</td>
<td>has a sifter</td>
</tr>
<tr>
<td>made of tin</td>
<td>made of tin</td>
</tr>
<tr>
<td>powered by diesel fuel</td>
<td>powered by diesel fuel</td>
</tr>
<tr>
<td>operates on sea floor</td>
<td>operates on sea floor</td>
</tr>
<tr>
<td>works to remove lost fishing nets</td>
<td>works to remove broken glass</td>
</tr>
<tr>
<td>has a hook</td>
<td>has a metal pole with a sharpened end</td>
</tr>
<tr>
<td>made of tungsten</td>
<td>made of tungsten</td>
</tr>
<tr>
<td>powered by butane</td>
<td>powered by butane</td>
</tr>
<tr>
<td>operates on the beach</td>
<td>operates on the beach</td>
</tr>
<tr>
<td>works to remove broken glass</td>
<td>works to gather discarded paper</td>
</tr>
<tr>
<td>has a sifter</td>
<td>has a hook</td>
</tr>
<tr>
<td>made of brass</td>
<td>made of brass</td>
</tr>
<tr>
<td>powered by kerosene</td>
<td>powered by kerosene</td>
</tr>
</tbody>
</table>
not clearly plausible or implausible, it was excluded from the analyses presented below.

The 12 novel transfer items of Table 5 were printed on 3 x 5-in. index cards, each with five features, in the consistent order of Dimension A, B, C, D, then E.

Procedure. Classification learning and the two-feature classification test proceeded as in Experiment 2; that is, the contrast category was given the negatively defined label nonmorkel. Participants were then asked to classify each of the 12 novel transfer items as a morkel or a nonmorkel and to rate their confidence in this decision on a scale from 1 (guess) to 7 (certain). The order of presentation of the 12 novel transfer items was randomized for each participant.

Results

Classification learning. As expected, the classification-learning results largely replicate those found in Experiment 2. Eleven of the 16 coherent participants reached the learning criterion of two errors per block in 4 blocks or less, in comparison to 5 of the 16 incoherent participants. Under the assumption that nonlearners would reach criterion on the very next (i.e., fifth) block, the mean number of blocks to criterion was 3.25 and 4.60 for the coherent and incoherent groups, respectively, \( t(30) = 3.54, p < .005 \). The average error rate during learning for the coherent group (20.7) was reliably less than that for the incoherent group (30.5), \( t(30) = 2.84, p < .05 \).

Two-feature classification test. The two-feature classification test results also largely replicate those found in Experiment 2. The coherent group committed an average of 2.5 errors on the 18 two-feature items and achieved an average adjusted confidence score of 75.1, compared with scores of 8.0 and 12.5, respectively, for the incoherent group; \( t(30) = 7.10, p < .0001 \) for errors, and \( t(30) = 6.96, p < .0001 \) for confidence scores.

Even when only the 11 coherent and 5 incoherent participants who reached the learning criterion are considered, significant differences between groups remain on both error rates (1.5 for the coherent group vs. 7.4 for the incoherent group), \( t(14) = 6.40, p < .0001 \), and adjusted confidence scores (88.3 vs. 22.8), \( t(14) = 5.69, p < .0001 \).

Novel transfer test. The purpose of the novel transfer test in this experiment was to probe the level of abstraction of the morkel category induced by participants in the coherent condition. To this end, we classified participants’ responses as correct if they categorized a coherent novel machine as a morkel or an incoherent machine as a nonmorkel, and incorrect otherwise. In fact, the coherent group committed significantly fewer errors on the novel transfer items than the incoherent group, 2.5 versus 4.9, respectively, out of a possible 11, \( t(30) = 2.47, p < .05 \).

Just as for the two-feature classification test, an adjusted confidence score was calculated by negating the confidence ratings on those novel transfer items in which the classification decision was incorrect, and then summing the scores for each participant. The average adjusted confidence scores were 31.5 and 6.8 for the coherent and incoherent groups, respectively, a statistically reliable difference, \( t(30) = 2.68, p < .05 \).

These differences between the coherent and incoherent groups were not merely due to the greater number of learners in the coherent condition, as significant group differences remain even when only the 11 coherent and 5 incoherent learners are considered: As compared with the incoherent learners, coherent learners committed fewer errors (1.7 for the coherent group vs. 3.8 for the incoherent group), \( t(14) = 2.43, p < .05 \), and achieved a higher adjusted confidence score (38.8 vs. 18.4), \( t(14) = 2.25, p < .05 \).

Although the incoherent learners were less likely to classify the novel pollution-cleaning devices as morkels than the coherent learners, their performance on the novel transfer items was nevertheless significantly better than chance (3.8 errors vs. chance of 5.5 errors), \( t(4) = 4.54, p < .05 \); adjusted confidence score of 18.4 versus chance of 0, \( t(4) = 3.68, p < .05 \). That is, even though the morkel category they had learned was incoherent, they were still more likely to classify coherent novel items as morkels than novel incoherent items. There was no such tendency exhibited by the incoherent nonlearners, whose error rates and adjusted confidence scores on the novel transfer test (5.6 and 2.9, respectively) did not differ from chance responding.

Discussion

The purpose of Experiment 4 was to test whether participants in the coherent group represented morkels as a disjunctive category in which category membership was predicted by the water morkel or the land morkel or the air morkel. In fact, the results demonstrated that the category representation learned by the coherent group supported a form of generalization, or abstraction, beyond the three concrete morkel subtypes, because novel devices that possessed none of the features of the original three morkels were considered morkels when the particular combination of novel features supported an interpretation of the device as a pollution-cleaning machine. This result cannot be attributed either to the larger number of features shared by the coherent novel machines and the original three morkels or to the greater semantic similarity of their features, because the novel coherent and incoherent machines were constructed from the same set of features.

Unexpectedly, participants in the incoherent condition who learned were also more likely to classify coherent novel machines as morkels than incoherent novel machines. This is a surprise, because one might have expected that these participants would have been more likely to classify the incoherent novel machines as morkels given the lack of coherence exhibited by the three original morkels they learned. We suggest two explanations for this result. First, incoherent learners may have been manifesting a general expectation that labeled objects would designate coherent objects, or a more specific expectation that morkel designated any kind of coherent pollution-cleaning device, and they may have maintained these expectations despite the observation of incoherent morkels during learning. Indeed, our analysis of the debriefing questions from the previous experiments indicated that most incoherent learners had concluded that morkels must be some kind of pollution-cleaning device. This account also explains why incoherent nonlearners (who were somewhat less likely to describe incoherent morkels as pollution-cleaning devices) were less likely to classify coherent novel machines as morkels. A second explanation is that the incoherent learners may have found a way to make sense of the incoherent morkels, and as a result were more likely to classify coherent versus incoherent novel devices as morkels.

Nevertheless, despite the abstract knowledge demonstrated by incoherent participants, participants in the coherent condition were significantly more likely to classify the coherent novel machines as
morkels. This result cannot be attributed to the greater number of features shared by morkels and the novel coherent machines in the coherent versus incoherent conditions or to those features’ greater semantic similarities, because the coherent and incoherent morkel categories were constructed from the same features. Instead, we attribute this difference to the relational structure shared by the original morkels and the novel coherent machines in the coherent but not the incoherent condition. That is, coherent participants induced an abstract category representation that consisted of a system of relations, specifically, that the gathering instrument is suitable for collecting the type of pollution and that the location be an area where the pollution might be found.

One might question to what extent membership in the coherent morkel category in fact depended on satisfaction of these specific two relations. One possibility is that more coherent than incoherent participants induced that morkels were a pollution-cleaning device, and, as a result, morkel category membership was simply determined by the extent to which the novel devices could be construed as coherent pollution-cleaning devices, regardless of whether they possessed a (coherent) gathering instrument or operating location. Another possibility is that the group difference arose because incoherent learners were somewhat influenced by the incoherent category exemplars they observed previously. However, our analysis of the debriefing questions indicated that incoherent learners were more likely to describe morkels as “pollution-cleaning devices” (although this difference was not statistically significant), and very few described morkels as “machines that don’t make sense.” Finally, the group difference did not arise from a minority of incoherent learners being more likely to classify the incoherent novel machines as morkels, because all 5 of the incoherent learners were more likely to classify the coherent versus incoherent novel machines as morkels.

General Discussion

The experiments reported in this article demonstrate that the system of interfature relationships shared by category members had two important effects on the acquisition of a novel category. First, in four experiments, participants learned the coherent category more readily than the incoherent category. Because these two categories were constructed from the same features and had the same empirical structure, this result must be attributed to the relations that interrelate the features in the coherent condition. Second, learners of the coherent category formed an abstract category representation that led them to categorize new machines as morkels, even when they shared none of the features of the original category members. In the following sections, we first discuss the learning results and then the nature of the morkel category abstraction.

The Learning of Abstract Coherent Categories

The learning advantage exhibited by the coherent group was obtained across a number of learning conditions, conditions that varied the extent to which they focused participants’ attention on within-category attribute relations. When predicting a missing attribute value from an exemplar’s other features (Experiment 1), the average coherent participant made only one error during learning, and 3 (of the 12) coherent participants made no errors during learning at all. These results stand in stark contrast to the learning usually exhibited in categorization research, which typically requires either extensive feedback, or at least extensive exposure to stimulus materials. Instead, the nature of learning that took place is perhaps best described as an instance of learning-by-comprehension, in which a stimulus spontaneously elicits and combines with the knowledge that the learner already possesses to construct a new cognitive representation. That learning in fact took place (i.e., that the novel representation was eventually associated with the new category label morkel) was evidenced by the subsequent good performance of the coherent group on the 18 two-feature transfer items.

Learning was more difficult in Experiments 2, 3, and 4 when participants learned to discriminate coherent morkels from members of an incoherent contrast category. We attributed this greater difficulty to participants searching for simple between-category differences at the expense of those within-category attribute correlations required for successful learning. For example, the delayed learning (as compared with Experiment 1) may have been due to participants first searching for single-dimension rules with which to discriminate the categories, and, when this strategy failed, only noticing the meaningful interattribute morkel relations when searching for conjunctive rules (Nosofsky, Palmeri, & McKinley, 1994). Alternatively, participants may have first formed an (incorrect) prototype representation of the morkel category that encoded individual feature frequencies but not interfature correlations, and, when this strategy failed, only apprehended the morkel subtypes when encoding individual exemplars (Smith & Minda, 1998).

Nevertheless, despite the delayed learning experienced by the coherent groups in Experiments 2, 3, and 4, relative to Experiment 1, those participants still demonstrated faster learning compared with those in the incoherent conditions. Evidence that this result was due to coherent participants apprehending the meaningful morkel subtypes is provided by the classification on two-feature transfer items, which showed that the coherent but not the incoherent learners encoded virtually all the interfature correlational information exhibited by the observed category members, and by our analysis of the debriefing questions. This result also indicates that these groups differed not just on how fast they learned but also on what they learned.

Although their learning proceeded more slowly than the coherent participants, some participants in the incoherent condition achieved success predicting missing features (in Experiment 1) and discriminating morkels from nonmorkels (in Experiments 2 and 4), and did so despite the absence of meaningful relations linking category features. This success is readily explained in terms of existing rule-based and similarity-based models of category learning. For example, using categories exhibiting highly intercorrelated features like the morkel category, Billman and Knutson (1996) showed that participants exhibit learning of those correlations even during an unsupervised learning procedure, a result they attribute to the acquisition of rules that predict one feature from another (also see Billman & Heit, 1988; Kersten & Billman, 1997). Documenting how people can learn to discriminate category membership on the basis of similarity to category exemplars (Medin & Schaffer, 1978; Nosofsky, 1986) or rules (Bourne, 1970; Nosofsky et al., 1994) has constituted the bulk of categorization research for decades. These theories also explain the
difficulty of learning to discriminate morks from nonmorks in Experiments 2 and 4, because each predicts the difficulty of learning categories that completely overlap in their constituent features (because of the high degree of between-category similarity; Rosch & Mervis, 1975; or because of people's tendency to first look for single-dimension rules first; Nosofsky et al., 1994). Conceivably, the difficulty of learning may also have been due to the incoherent participants wasting cognitive resources trying to make sense of the verbal descriptions of the incoherent morks.

The presence of a label for the contrast category in Experiment 3 (krenshaws) led to a lower percentage of participants reaching the learning criterion as compared with when no label was provided in Experiments 2 and 4 for both the coherent (42% vs. 72%) and the incoherent (0% vs. 29%) groups. We suggest three possible reasons for this result. First, the contrast category's positive label may have led participants to test additional (incorrect) single-dimension rules associated with the contrast category before considering conjunctive rules. Second, the additional label may have induced participants to form and use an (incorrect) prototype representation of the contrast category (Goldstone, 1996). Finally, the symmetrical category labels may have induced both coherent and incoherent participants to waste cognitive resources looking for meaningful interattribute relationships in the contrast category where none was to be found.

Our results revealed that participants in both the coherent and incoherent groups also learned something about the morkel categories that we have characterized as abstract, because it includes information that goes beyond the information encoded in any one category exemplar. For example, our analysis of the debriefing questions indicated that both groups induced that morks were some kind of pollution-cleaning device. Of course, the detailed nature of the learning processes that produce such category-level abstractions is an open question of considerable theoretical importance. One possibility is that the abstractions formed in the current experiments arose from participants comparing the individual coherent morks and abstracting their common components. Indeed, previous research has shown that category learners will focus on features shared by those category exemplars that they compare (e.g., as a result of reminders among category members, see Medin & Florian, 1992; Ross & Spalding, 1991; Spalding & Ross, 1994). In the current experiments, participants' understanding of morks as some kind of pollution-cleaning device was likely to have resulted from them comparing and abstracting over the features of individual morks such as "works on absorbing spilled oil," "works to gather harmful solids," and "works to absorb dangerous gaseous ions."

In addition to inducing that morks were pollution-cleaning devices, a central claim of the current article is that participants in the coherent condition also induced the abstract system of relations that characterizes morks, namely, that they possess a collecting instrument and an operating location suitable to the type of pollution they gather. This conclusion was based on the results of Experiment 4's novel transfer test in which coherent participants were more likely than incoherent participants to classify coherent novel machines as morks. We suggest that this result was also likely to have arisen from coherent participants comparing the three original morkel examples and extracting the interfeature relations they had in common (the means of gathering and the device's operating location are appropriate to the type of pollution; also see Lassaline & Murphy, 1998; Ramscar & Pain, 1996; and see Kuehne, Forbus, Gentner, & Quinn, 2000, for a possible computational model of such an abstraction process). This abstract relational structure then contributed to the degree of morkel category membership for those novel machines that also displayed that structure. We discuss the means by which categorization in the presence of abstract relations is carried out in the following section.

One potential concern with the current study might be whether the morkel category was truly novel. That is, it is conceivable that some participants had a prior concept of pollution-cleaning device and that the learning that took place was to merely associate an existing category with a new label (morkels). Although our own intuitions find this possibility unlikely, we note that this prior conception of pollution-cleaning devices fails to explain why coherent participants were more likely than incoherent participants to classify coherent novel machines as morks. That is, the common relational structure observed by the coherent participants in the original morkel category members contributed to the degree of morkel category membership as compared with the incoherent condition, even if both groups had some prior conception of pollution-cleaning devices. At a minimum, coherent participants must have constructed a new category representation that refined their existing notion of pollution-cleaning devices to include the specific relational structure embodied in the original exemplars of the category.

Categorization With Abstract Coherent Categories

It is illuminating to consider the characteristics of the categorization processes that led to certain novel machines being classified as morks in light of the coherent morkel's abstract relational structure. We suggest, that in the presence of a system of relations, category membership is determined by the extent to which a given combination of features displayed by a to-be-classified exemplar satisfies those relations; that is, the extent to which the concrete features instantiating the relational structure were mutually compatible with one another with respect to the relationships that linked them. Of importance, what it means for two features to be compatible with respect to a relationship is not encoded as part of the category representation but rather depends on general background knowledge. For example, the novel machine that "operates in highway tunnels," "works to remove carbon monoxide," and "has large intake fans" was considered a morkel because fans might be good at removing carbon monoxide (the type of pollution is compatible with the gathering instrument), and carbon monoxide is likely to be found in highway tunnels (the location is compatible with the type of pollution). It should be noted that this determination could only have been made in light of background knowledge about carbon monoxide, namely, that it is a gas (and hence might be removed with a fan) emitted by automobiles (and hence might be found in highway tunnels). In contrast, the novel machine that "operates in highway tunnels," "works to remove malaria-ridden mosquitoes," and "has a large magnet" was not considered a morkel, because background knowledge informs us that malaria-ridden mosquitoes are unlikely to be found in highway tunnels or gathered with a magnet. Thus, background knowledge is not only crucial in apprehending the original three coherent morks and inducing the abstract relational structure of the category, it is also
critical in determining whether the features of new to-be-classified machines satisfy that relational structure.

The emphasis on the role of relations among abstract features in determining category membership rather than on concrete features is the defining characteristic of abstract coherent categories. This role of relations distinguishes abstract coherent categories from the types of categories addressed by the well-known similarity-based models of categorization (Medin & Schaffer, 1978; Nosofsky, 1986; Rosch & Mervis, 1975). In such models, mental conceptions of categories consist of, for example, a central prototype or previously stored category exemplars, and category membership is a function of the overall similarity of a new to-be-classified object to the prototype (or the exemplars). These models achieve a degree of generalization, or abstraction, by assuming the existence of primitive within-dimension similarity relations, and that overall similarity is a function of the similarities on the individual dimensions. As a result, an exemplar exhibiting previously unobserved dimension values might be considered a category member if those values are similar to the corresponding values in the prototype or stored exemplars. However, primitive semantic similarity of features cannot explain the novel transfer results from Experiment 4, because the novel coherent and incoherent machines were constructed from the same set of features. Because of this identity of features across the two conditions, membership in the abstract coherent morkel category must have been a function of interactions among features rather than individual values on dimensions. We claim that these interactions consist of the satisfaction of interfature relations in light of background knowledge.

Categories defined by relations rather than concrete features have the quality that their extension becomes virtually infinite, because there are a virtually infinite number of feature combinations that might satisfy the relational structure. For example, we trust the reader would agree that the six novel morkels shown in the left-hand side of Table 5 do not begin to exhaust the collections of features that might be construed as coherent pollution-cleaning devices, and hence referred to as morkels. The fact that a future machine that cleans up a new type of pollution (with some suitable gathering instrument in some suitable location) might come to be called a morkel also illustrates the open-ended extension of the category. One should note that, technically speaking, categories defined in terms of similarity may also possess an infinite extension, because when one or more of the stimulus dimensions is continuous, there exists an infinite number of exemplars that are highly similar to the prototype (or one of the exemplars). However, such categories are defined with respect to a stimulus space with a fixed set of concrete dimensions, dimensions determined by the original exemplars of the category. In contrast, the role of background knowledge in interpreting whether concrete features satisfy an abstract coherent category’s interfature relationships means that such categories are liberated from any predefined stimulus space. For example, after the coherent morkel category was learned, we doubt that the feature “has a metal pole with a sharpened end” occupied some position in a multidimensional morkel stimulus space, because sharpened metal poles are not normally thought of as gathering instruments. Rather, a sharpened metal pole became interpretable as a gathering instrument only in the context of the additional features “works to gather discarded paper” and “operates in parks.” Our use of the description “open-ended” is intended to reflect the presence of categorization processes that provide flexible interpretations of concrete features in light of the relationships that interconnect them.

The human conceptual system appears to be heavily populated with concepts that are abstract and coherent in the ways we have been discussing. For example, political revolution consists of an adversarial relationship between two parties, one possessing political power and attempting to retain it, the other lacking such power and attempting to acquire it. Military invasion and political subversion similarly consist of particular types of adversarial relationships between particular types of groups (Abelson, 1981). Other social groupings include families (a particular system of relations between the abstract features “father,” “mother,” “children”), corporations (board of directors, president, departments, employees), governments (legislature, judiciary), and academic departments (faculty, courses). Legal concepts include murder (the killer and the killed), robbery (the robber, the robbed, and the stolen property), marriage and divorce (husband and wife), and lawsuit (the damaged, damages). In each case, the category representation consists of relations between abstract features that constrain how those abstract features may be jointly instantiated. For example, in the case of political revolution, the two parties are constrained to belong to the same country. In the case of (traditional) families, the children are constrained to be of the same race as their parents. Finally, many superordinate categories appear to be abstract and coherent, because the basic-level categories below them in a taxonomic hierarchy have been described as the highest-level categories associated with a concrete image (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) and because many superordinate categories also possess the type of relational structures we have been describing. For example, the superordinate category mammal appears to possess a rich relational structure that encodes knowledge such as the fact that mammals bear live offspring, the females possess mammary glands, and so on.

Language may be the source for many abstract coherent categories. For example, it is unlikely that many of us have observed enough political revolutions to have induced the concept from observations alone. Instead, some abstract coherent categories may be obtained by definition or induced from the linguistic context in which the category label appears (Landauer & Dumais, 1997). Nevertheless, a central contribution of the current article is to demonstrate that abstract coherent categories can also be induced from observations when the appropriate feedback is supplied.

The sensitivity to relations that we have argued distinguishes the abstract category learned by coherent versus incoherent participants in the current study also distinguishes the category learned by the coherent group from those learned in some other studies of category learning. For example, although in Wisniewski and Medin’s (1994) study, participants used abstract concepts such as detailed, detailed does not seem to possess an internal structure consisting of multiple abstract features interconnected by relations. Likewise, the category “things to take out of one’s house in case of fire” is also abstract (category members share few concrete features) and perhaps has several abstract features (“valuable,” “can be removed from house quickly”), but there do not appear to be relations among the abstract features that constrain how they may be jointly instantiated (although the fact that bricks of gold but not babies can be thrown from burning windows might be a counterexample). Finally, although we have suggested that many superordinate categories are not only abstract but also coherent,
this may not always hold; for example, there do not appear to be strong constraints on how the abstract features of "furniture," such as "manufactured," "functional," and "found in living spaces," are jointly instantiated.

There is still more to be learned about abstract coherent categories. For example, because our conclusions regarding the morcell category abstraction rests in part on our analysis of the debriefing questions, future research that more directly assesses the scope of the generalization supported by abstract coherent categories would be of value. Extending the technique introduced in this article by contrasting additional types of novel transfer items (e.g., the current novel items vs. novel coherent pollution-cleaning devices that don’t specify a collecting instrument or location vs. coherent devices that don’t gather pollution) is one way to carry out such assessments.

**Conclusion**

The current research contributes to the body of evidence that emphasizes the knowledge that interrelates the features of human category concepts and which makes our conception of those categories meaningful and coherent. It demonstrates that the role of relationships can go beyond merely making certain combinations of known features acceptable for category membership. Rather, the system of relations provide an interpretive framework in which combinations of completely novel features may be viewed as acceptable category members. These abstract coherent concepts appear to be part and parcel of much of our everyday thought, including conceptions of social groups, legal concepts, societal institutions, political and military scenarios, and classes of natural kinds.

**References**


Received February 1, 2000
Revision received December 11, 2000
Accepted January 2, 2001