ELSEVIER

Cognition 69 (1998) 135–178

COGNITION

Why are different features central for natural kinds and artifacts?: the role of causal status in determining feature centrality

Woo-kyoung Ahn*

Department of Psychology, Yale University, 2 Hillhouse Avenue, P.O. Box 208205, New Haven, CT 06520-8205, USA

Received 6 May 1998; accepted 14 September 1998

Abstract

Ahn and Lassaline [Ahn, W., Lassaline, M.E., 1995. Causal structure in categorization. Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society, Pittsburgh, PA, pp. 521–526] recently proposed a causal status hypothesis which states that features that play a causal role in a relational structure are more central than their effects. This hypothesis can account for previous research demonstrating that compositional features are generally important for natural kinds but functional features are generally important for artifacts. The causal status hypothesis explains this category-feature interaction effect in terms of differences in the causal status of compositional and functional features between natural kinds and artifacts. Experiments 1 and 2 examined real-life categories used in previous studies, and found positive correlations between the causal status of the features and their centrality across natural and artifactual kinds. Experiments 3 and 4 manipulated the causal status of compositional and functional features in artificial categories, and showed that it was causal status rather than the interaction between the type of feature and the type of category per se that accounted for feature centrality. The implications of these results on the distinctions between natural kinds and artifacts are discussed.

1. Introduction

One traditional assumption in categorization research is that there must be a categorization process common to various types of concepts, and that this process

* Fax: +1 203 4327172; e-mail: ahn@pantheon.yale.edu

0010-0277/98/$19.00
PII: S0010-0277(98)00063-8

can be studied using generic experimental stimuli (e.g. Bruner et al., 1978; Anderson, 1991; Nosofsky et al., 1994). Although such a domain-general approach can sometimes be useful, an increasing number of studies have demonstrated the effect of domain-specific knowledge in categorization (e.g. Carey, 1985; Murphy and Medin, 1985; Pazzani, 1991; Ahn et al., 1992; Wattenmaker, 1995). As a result, the growing view in categorization is that it is difficult to assume a single domain-general, all-purpose categorization process spanning across diverse domains, such as natural kinds, social categories, and artifacts (see Hirschfeld and Gelman, 1994 for a review).

Of particular concern in this paper is the phenomenon that different kinds of features are shown to be important for different kinds of categories. Many studies (e.g. Gelman, 1988; Barton and Komatsu, 1989; Keil, 1989; Rips, 1989) have shown that feature centrality depends both on the kind of category to which the object belongs and on the type of feature. In general, features internal to the object, such as molecular structure, were considered more important for categorization of an object as a natural kind than for categorization of an object as an artifact. In contrast, features external to the object, such as function, were generally shown to be more important for artifact category membership than for natural kind category membership (but also see Malt and Johnson, 1992, which will be discussed later). Why are different types of features central depending on whether a category is a natural kind or an artifact? This introduction first briefly reviews research demonstrating an interaction between category type and feature type, followed by a presentation of the causal status hypothesis (Ahn and Lassaline, 1995; Ahn, Lassaline and Kim, in preparation) which provides an account for this interaction.

1.1. Category-feature interaction

The interaction between the type of category and the type of feature has been observed with both adults and children. In the study of Gelman (1988) children first learned a novel feature for each type of category (e.g. this rabbit has a spleen inside) and were asked whether this feature is generalizable to another instance of the same category. The second graders in this study responded that features referring to substance and internal structure (e.g. 'has a spleen inside', p. 74) were more generalizable for natural kinds, whereas functional features (e.g. 'you can loll with it', p. 75) were more generalizable for artifacts.

Keil (1989) demonstrated a similar phenomenon using a transformation task. For example, children learned a story that a raccoon was dyed and painted with a white stripe, or a story that a coffeepot was transformed into a bird feeder by changing its parts. Fourth graders as well as adults in this study judged that changes in perceptual appearance did not matter for natural kinds' identity, but these changes did matter for artifacts' identity. Presumably, a perceptual change in an artifact is directly related to the function it performs, and, therefore, perceptual changes were judged to matter for artifact category membership. However, a new discovery about the origin of an artifact (e.g. a key that was, in fact, made of pennies) did not affect category membership.

Similarly, Rips (1989) asked adult participants to read stories about various kinds of transformations. In one story, an animal called a sorp transformed from a bird-like creature into an insect-like creature. In this story, the transformations were either due to chemical hazards or the result of maturation. When the transformation was made externally (i.e. chemical hazards), it was less likely to change the natural kind's identity than when it was made internally (i.e. maturation). With artifacts, however, the critical determinant in changing their membership was shown to be the change in the function that the designer of the artifact intended to produce.

Barton and Komatsu (1989) also presented adult subjects with either natural kinds (e.g. goat) or artifacts (e.g. tire) with different kinds of change introduced. The changes included what they termed molecular changes (e.g. a goat having a changed chromosomal structure or a tire not made of rubber) and functional changes (e.g. a female goat not giving milk or a tire that cannot roll). The results showed that the molecular changes mattered more for natural kind membership whereas the functional changes mattered more for artifact membership.

Barr and Caplan (1987) examined what they termed 'intrinsic' and 'extrinsic' features. A feature is intrinsic if it can be true of an entity considered in isolation (e.g. 'has wings' in a bird). A feature is extrinsic if it is represented as the relationship between two or more entities (e.g. 'used to work with' for a hammer). According to these definitions, the molecular structure, substance, or internal features of an object would be intrinsic features whereas functional features would be extrinsic features. Consistent with the above studies, Barr and Caplan (1987) (Experiment 3) also found that artifact categories such as weapons and toys are more likely to be defined in terms of extrinsic features than natural kinds such as mammals and trees.

An immediate conclusion that one might be tempted to draw from this interaction effect is that natural kinds and artifacts are two separate domains and that different features are more central in categorization as a result of this domain distinction. Barton and Komatsu (1989) advocate the most extreme version of this conclusion by stating that 'our results suggest that having a particular molecular or chromosomal structure is necessary and sufficient for (i.e. defining of) membership in a particular natural kind category and that having a particular function is necessary and, perhaps to a lesser extent, sufficient for membership in a particular artifact category (p. 444).'

However, this conclusion seems invalid for a number of reasons. First of all, they assume that natural kinds and artifacts have defining features. To the contrary, Hampton (1995) has failed to identify clearly necessary features for a range of concepts including both artifacts and natural kinds. Second, Barton and Komatsu's conclusion seems to be problematic even if it is weakened so that the claim is that the weights of features, rather than which features are defining, vary depending on the domain. The weaker version would be that molecular features in natural kinds and functional features in artifacts might not be necessary and sufficient, but they are more important than other features. However, this weaker version still does not fare well, because Malt and Johnson (1992) have demonstrated that physical properties of artifacts (e.g. for boats: 'is wedge-shaped, with a sail, anchor, and wooden sides', as selected by undergraduate subjects in the study) were judged to be more important

than, or just as important as, functional features (e.g. 'to carry one or more people over a body of water for purposes of work or recreation').

An alternative way of explaining the category-feature interaction can be suggested using the so-called theory-based approach to categorization (Carey, 1985; Murphy and Medin, 1985; Keil, 1989; see the special issue of Cognition (1998, 67/1,2) for a thorough review). The idea is that conceptual representation is analogous to theory representation in that they are comprised of richly structured features rather than a set of unrelated or independent features. The theory-based approach also argues that concepts are embedded in one's domain theory and the centrality of features is determined by their importance in these principles or theories underlying the categories (Murphy and Medin, 1985). For example, in Keil (1989), although older children believed that a raccoon transformed into a skunk at the perceptual level is still a raccoon, younger children (i.e. kindergartners) tend to draw their conclusion based on perceptual appearance. This result was presumably obtained because kindergartners did not yet have well-developed causal theories about the properties.

In the previous theory-based approach, however, the exact mechanism of how domain theories operate in categorization has rarely been articulated in detail (Gelman and Kalish, 1993; Murphy, 1993). Most relevant to the current study is that it is not clear exactly what it means for a feature to play a central role in one's domain theory. For instance, in the study of Malt and Johnson (1992) mentioned earlier, physical features were judged to be more important than functional ones in artifacts. The theory-based approach might say that this is because the physical features in those artifacts serve a central position in the domain theories of the artifacts, but such an account sounds post-hoc in the absence of an independent way of determining how a feature plays a central role in domain theories.

Thus, the purpose of the current study is to introduce a more precisely defined theory-based account for feature centrality called a causal status hypothesis, and to apply this hypothesis to explain the category-feature interaction effect as well as the absence of the effect. The next section first explains the causal status hypothesis, and then attempts to explain the existing data with this hypothesis.

1.2. Causal status hypothesis

The traditional view of concepts is that concepts are clusters of correlated features (Rosch, 1978; Medin et al., 1982; Malt and Smith, 1984; Billman, 1989; Billman and Knutson, 1996). Murphy and Medin (1985) took this one step further and argued that people deduce reasons underlying these correlations. Hence, a connection between features is not just a simple link but rather 'a whole causal explanation for how the two are related (Murphy and Medin, 1985, p. 300)'. Similarly, other researchers who study the role of theories in categorization have generally accepted causality as a central component in theory-like conceptual representation (Carey, 1985; Wellman, 1990; Gelman and Kalish, 1993). Given a set of causally related features, the causal status hypothesis states that people have a bias toward weighting features that serve as causes of other features more than their effects.

It is easiest to illustrate the causal status hypothesis by explaining the paradigm employed in Ahn et al. (in preparation). Participants in their Experiment 1 first received descriptions of three characteristic features of hypothetical kinds. For example, they learned that if a person has features A, B, and C[1], 75% of the time they have Disease Xeno. Participants were also told that symptom A causes symptom B, which in turn causes symptom C. Then, they received descriptions of objects in which one of the three features is missing. The missing feature was either the most fundamental cause (symptom A), the intermediate cause (symptom B), or the terminal effect (symptom C). When asked to estimate the likelihood that each object belonged to the target category, participants judged the item missing the most fundamental cause to be least likely, and the item missing the most terminal effect to be most likely. The item missing the intermediate feature in the causal chain was placed right in the middle, being statistically different from the features at both ends. Hence, the results strongly supported the causal status hypothesis; the more causal a feature was, the more influence it had on categorization.

Other follow-up experiments used more complex causal structures involving multiple causes or multiple effects (Experiment 2, Ahn et al., in preparation). For instance, depression might have multiple causes such as a low dopamine level or the loss of loved one. In addition, a low dopamine level in the brain might lead to multiple effects such as depression or schizophrenia. When presented with an option for categorizing objects either by cause or by effect, participants preferred categorizing objects based on matching cause rather than matching effect. For instance, a person who is depressed all the time (i.e. effect feature) because she has low levels of dopamine (i.e. cause feature) was more likely to be categorized with a person matching in cause (a person who suffers from insomnia because she has low levels of dopamine) than with a person matching in effect (a person who is depressed all the time because she has been taking alpha-methyldopa). Hence, the causal status effect was generalized to situations involving multiple causes and effects.

Why would people value cause features more than effect features in their conceptual representation? Concepts organized around causes seem to provide more inductive power than those organized around their effect. For instance, if a physician's disease concepts are organized based on the surface symptoms, for example, diseases that all involve weight loss, such concepts would lack both predictability and controllability. Without an understanding of the causal mechanisms underlying the symptoms, it would be difficult or impossible for a doctor to predict the progress of the disease, let alone prescribe a treatment plan. Likewise, social concepts would be more useful if they were organized by causal features (e.g. personality traits) rather than by surface features (e.g. behavioral characteristics) because one could then predict other behaviors in completely novel situations based on the cause. For instance, a person might be an extremely cautious driver, but it would be difficult to generalize that she would be also very cautious in her work because it would depend on why she is careful at driving (e.g. is she in general careful, does she always carry fragile objects in her car, etc.). As such, it is speculated that one of the

[1]These are not the actual features used in this study, but are used here for ease of understanding.

reasons why people weigh cause features more than effect features has to do with the generative power of cause features in inductive reasoning.

In addition, features that are more causal seem more immutable in conceptual representations. Sloman et al. (1998) defined mutability of features as the ease with which people can transform their mental representation of an object by eliminating a feature or by replacing it with a different value, without changing other aspects of the object's representation. The more causal a feature is, the more difficult it seems to mutate the feature without changing other aspects of the conceptual representation. For instance, if we are to imagine a new breed of dog that does not have a hippocampus (which presumably causes many behaviors of dogs), we need to alter a lot of features in our dog concept including their behavior and even their status as a pet. On the other hand, imagining a new kind of dog that does not have a tail would not require much conceptual mutation except that they would not wag their tails. As such, more causally central features seem more responsible for conceptual coherence and consequently would be judged to be more central in categorization.

Before preceding to the main focus of the current study, it is necessary to specify other important assumptions of the causal status hypothesis. First, causal relations in the causal status hypothesis are used in a broad sense in that they include not only physical causal relations (e.g. A's movement caused B's movement), but also other explanatory relations such as 'A enables B', 'A allows B', 'A determines B', 'A is a reason for B', 'B because of A', and 'A explains B'. This approach is in line with the theory-based approach, which serves as a basis for the causal status hypothesis. The idea is that conceptual representations are like theory representations, and as proposed by Carey (1985), 'Explanation is at the core of theories' (p. 201). It is beyond the scope of this paper to examine whether all of the above explanatory relations would lead to the causal status effect to the same degree in the same manner.

Second, the causal status hypothesis should not be taken as a claim that the causal status of a feature is the only determinant of feature centrality. Other possible determinants include cue validity, category validity, and perceptual saliency (see Sloman et al., 1998 for a review). In theory, all of these are orthogonal to each other in that one can manipulate one factor without changing other ones. Furthermore, the effect of one factor can be overridden if there is enough counteracting force (e.g. effect feature has a high cue validity). Therefore, in making a prediction of the causal status hypothesis, it is important to equate the other factors. That is, the cause feature would be considered more central than its effect, all else being equal.

1.3. Account for domain differences in feature weighting

The purpose of the current study is to demonstrate that the interaction effect between the type of category (natural kinds versus artifacts) and the type of feature (molecules versus functions) occurs because of differences in the causal status of features. Consider the study of Barton and Komatsu (1989) again. The molecular features in their natural kinds (and probably in most real-life natural kinds) tend to

serve as causes of other features, including functional properties and physical appearance. For example, a goat's DNA structure enables milk production, which, according to the causal status hypothesis, is the reason why molecular features were judged to be more central in natural kinds. In artifacts, however, features seem causally structured around their functional rather than molecular features. Going back to Rips' study (Rips, 1989), intended functions (i.e. the function intended by the designer who produced the artifact) had a larger impact on artifact membership than substance. According to the causal status hypothesis, this occurred because intended functions determine what substance should be used in making the object. For example, because a person intends to make an object used for sitting, a chair is made of wood or metal rather than tofu or jelly.

In the study of Malt and Johnson (1992), however, physical features were judged to be more important than functional features. These results seem to have occurred because sometimes substance or physical appearance in artifacts allows the artifacts to perform certain functions. For instance, sweaters provide warmth because they are made of wool. In that case, 'being made of wool' (physical feature) would be judged to be central. Note, however, that there are some physical features of artifacts that are not functionally relevant or do not seem to cause any features for that matter. For instance, the color of computers tends to be taupe, but does not determine the computer's function or any other features, and therefore is predicted by the causal status hypothesis to be peripheral. Indeed, a closer examination of the stimulus materials of Barton and Komatsu (1989) suggests that the physical features they used for artifacts seemed causally irrelevant (e.g. box-like shape for TV, or cylinder shape for pencils). This difference in the causal status of physical features seems to explain the discrepancy between Malt and Johnson (1992) and Barton and Komatsu (1989).

The current claim clearly contrasts with the account given by Barton and Komatsu (1989) as well as its weaker version. According to the causal status hypothesis, molecular features, for example, might not be defining or central features for natural kinds if they do not serve as a cause for other features. Furthermore, even functional features might become less essential in artifacts if they serve as mere effects of other features. Conversely, functional features can be essential in natural kinds if they cause other features. For instance, 'giving milk' and 'producing fertile eggs' in goats would be judged to be more central if one's domain knowledge includes that these features also lead to raising and producing baby goats.

1.4. Overview of experiments

The four experiments reported here are divided into two parts. Part A examines real-life natural kinds and artifacts,. Since these are categories existing in the real world, the causal status of each feature was measured rather than manipulated, and then was correlated with the feature's centrality in a categorization task. Part B uses novel artifacts and natural kinds and manipulates the features' causal status in order to examine whether changes in the causal status of features can actually induce changes in the features' centrality.

2. Part A: examination of real-life categories

The task and logic behind Experiments 1 and 2 were similar. Participants carried out a categorization task and a causal judgment task in a counterbalanced order. In the categorization task, participants were presented with an object missing a feature, and were asked to judge the likelihood that the object belongs to a certain category. For example, participants in Experiment 1 judged the likelihood that an object would be a goat if it were in all ways like a goat except that it did not give milk. The more likely it was that this object was a member of the category, the less important the missing feature would be. The causal judgment task assessed the causal or explanatory relations among features used within each of the categories in the categorization task. In the case of a goat, participants in Experiment 1 were asked whether a goat can give milk because it has a 'goat' genetic code. The idea is that the ratings on the causal judgment task should predict ratings on the categorization task. That is, the more causal a feature is as measured in the causal judgment task, the more central it should be, as measured in the categorization task. In brief, Experiments 1 and 2 examined the correlation between the causal status of features in existing concepts and the centrality of these features. Experiment 1 used the stimulus materials of Barton and Komatsu (1989), which had well-distinguished types of features (molecular, physical appearance, and functional features). Experiment 2 used the materials of Malt and Johnson (1992), which had 12 artifacts with physical and functional features.

2.1. Experiment 1

As explained in the Introduction, Barton and Komatsu (1989) found an interaction effect between feature type and category type. According to the causal status hypothesis, this interaction effect is due to the fact that different types of features have different causal status depending on which category they belong to. For example, molecular features in natural kinds tend to cause their physical appearance and/or their functions, whereas in artifacts, molecular features are determined by the functions they are supposed to perform. The predictions of Experiment 1 are that the categorization task will replicate Barton and Komatsu (1989), and that the causal judgment ratings will correlate with the categorization ratings.

2.1.1. Participants

Seventeen undergraduate students at Yale University participated in this experiment in partial fulfillment of the requirements for an Introduction to Psychology course.

2.1.2. Materials and design

Table 1 shows the full set of features and categories broken down into each type. The feature types were molecular features, functional features, and physical appearance. The category types were natural kinds and artifacts.

Using these features and categories, the questions for the categorization task were

Table 1
Materials used in Experiment 1

Category	Functional feature	Molecular feature	Physical feature
Artifacts			
Mirror	1. Reflects an image	2. Glass	3. Hard
Pencil	4. Can write	5. Lead	6. Cylindrical
Record	7. Music when played	8. Plastic	9. Round
Tire	10. Rolls	11. Rubber	12. Filled with air
TV	13. Show visible pictures	14. Screen made of glass	15. Box-like
Natural kinds			
Goat	16. Gives milk	17. Genetic code	18. Has four legs
Gold	19. Used in dental work	20. Atomic structure	21. Yellow
Horse	22. Races	23. Genetic code	24. Has mane
Tree	25. Burns	26. Genetic code	27. Has a trunk
Water	28. Tastes good to drink	29. Atomic structure	30. Transparent

developed. For each of the features in each category, a question was developed in the form of 'Would X be still X if it were in all ways like X except that it did not have Y?' where X refers to a category and Y refers to a feature. This format was identical to the items used by Barton and Komatsu (1989). Along with each question, a 9-point scale was presented in which 9 indicated a rating of 'definitely yes' and 1 indicated a rating of 'definitely no'.

In addition, questions for the causal judgment task were developed for all possible pairs of features within each category. All questions took the form of 'category X has property A because of property B'. For example, a question for a record category involving its functional feature and its molecular feature was, 'Records are round because they are made of plastic'. For each pair, a 'reversed' question was also developed (e.g. 'Records are made of plastic because they are round'). Therefore, for each category there were six questions: molecular feature because of physical appearance, physical appearance because of molecular feature, molecular feature because of functional feature, functional feature because of molecular feature, functional feature because of physical appearance, and physical appearance because of functional feature. In total, there were 6×10 (the total number of categories) = 60 causal judgment questions.

2.1.3. Procedure

Ten participants first carried out the categorization task followed by the causal judgment task, and seven participants did the tasks in the reverse order. All experiments were conducted on PowerPC Macintosh computers. The experiment was programmed using Psyscope 1.1 (Cohen et al., 1993). All experiments were conducted in groups of 2–6 participants.

In the categorization task, participants were told that they would see a description of an object and be asked to judge the likelihood that an object belongs to a certain category. They further received the example, 'this thing is red, was originally designed to serve as a sign, but it cannot serve as a sign' along with a sample

question, 'How likely is it that this object is a stop sign?' They were instructed to answer the question by selecting a number on a 9-point scale on which 9 means very likely, 1 means very unlikely, and 5 means neither likely nor unlikely. Participants were told to respond by entering a number on the keyboard. They were informed that their responses would be displayed on the screen and that they could change their answers by using the delete key and then entering the correct answer. Once participants entered a response and hit the return key, which prompted the program to display the next question, they could not go back to their previous answers and correct them. Following the instructions and two practice questions, participants received 30 categorization questions in a randomized order.

In the causal judgment task, participants were instructed that for each problem they would see a sentence about an object in the form of 'A because of B' such as, 'a stop sign is painted in red because it is a visible color'. They were then instructed to judge how likely it was that the specified causal relationship was true. They were further instructed, using the above example, that if they thought this causal relation was definitely true, they should type in 9, and if they thought this causal relation was definitely false, they should type in 1. Participants were also instructed about entering their responses on the computer as in the categorization task. Following two practice questions, participants received 60 categorization questions in a randomized order. As in the categorization task, participants could correct their answers while working on a problem, but once they had proceeded to the next question, they could not go back and correct their previous answers.

2.1.4. Results

The pattern of the results was identical regardless of whether the participants carried out the categorization or the causal judgment task first. Therefore, the data were collapsed for the rest of the analyses. Fig. 1 shows the mean ratings from the two tasks for artifacts in the top panel and those for natural kinds in the bottom panel, along with the standard errors, as indicated by error bars. For ease of comparison between the two tasks, the ratings for the categorization task were reversed in the results section so that the higher the number is, the more important the feature is. The causal judgment data remained the same; the higher the number is, the more causal the feature is. The mean ratings for the causal judgment task used in the first part of the analysis were obtained as follows. There were four questions involving each feature, since there were three features in each category (e.g. for feature A, A because of B, A because of C, B because of A, and C because of A). Two of these questions concerned the likelihood that the feature served as a cause (e.g. for feature A, B because of A, and C because of A), and two other questions concerned the likelihood that the feature served as an effect (e.g. for feature A, A because of B, and A because of C). Fig. 1 reports the mean ratings of each feature serving as a cause.

2.1.5. Analysis of categorization ratings

As shown in Fig. 1, the categorization results replicated the results of Barton and Komatsu (1989). In artifacts, functional features (6.9) were much more important than molecular features (3.3) and physical appearance (3.0) whereas in natural

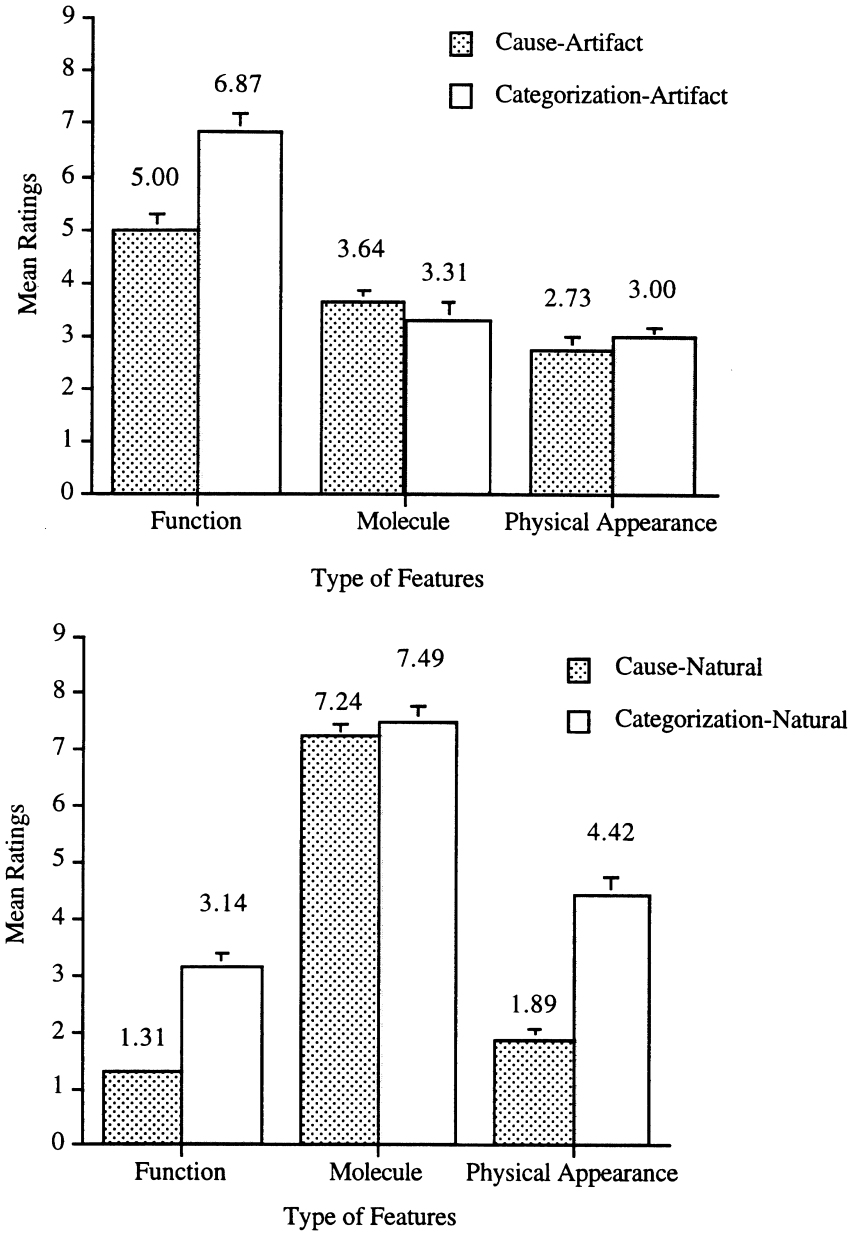

Fig. 1. Results from Experiment 1.

kinds, molecular features (7.5) were much more important than functional features (3.1) and physical appearance (4.4). An ANOVA with the type of feature and the type of category as within-subject variables supported this statement. The interac-

tion effect was reliable, $F(2,32) = 156.27$, MSe = 0.88, $P < 0.001$. In addition, there was a reliable main effect of the type of feature, $F(2,32) = 21.65$, MSe = 1.23, $P < 0.001$, because the ratings on physical features (3.7) were lower overall than the ratings on functional features (5.0) and molecular features (4.4). The main effect of type of category was not reliable, $P > 0.05$. Within natural kinds, pairwise comparisons indicated that mean ratings on all three features were highly reliably different from each other, $P < 0.0001$. Within artifacts, the ratings on the functional features were reliably different from those on both the molecular features and the physical features, $P < 0.001$, but the difference between the ratings on the molecular features and those on the physical features was not reliable, $P = 0.15$.

For an item analysis, a two-way ANOVA was conducted with the type of feature and the type of category as between-item variables. There was a reliable interaction effect, $F(2,24) = 43.94$, MSe = 40.11, $P < 0.001$, as well as a reliable main effect of the type of feature, $F(2,24) = 8.65$, MSe = 7.90, $P < 0.001$. There was no reliable main effect of the type of category.

2.1.6. Analysis of causal ratings

As in the categorization task, functional features in artifacts (5.0) were judged to be more causal than molecular features (3.6) and physical appearance (2.7), whereas in natural kinds, molecular features (7.2) were judged to be more causal than functional features (1.3) and physical appearance (1.9). An ANOVA with the type of feature and the type of category as within-subject variables supported this statement. The interaction effect between the type of feature and the type of category was reliable, $F(2,32) = 175.28$, MSe = 0.654, $P < 0.001$. There also was a reliable main effect of the type of feature, $F(2,32) = 216.59$, MSe = 0.41, $P < 0.001$, because the ratings on molecular features (5.4) were higher overall than the ratings on functional features (3.2), $P < 0.05$ which were also reliably higher than those on physical features (2.3), $P < 0.05$. There was no main effect of the type of category. Within natural kinds, pairwise comparisons indicated that all three features were reliably different from each other, $P < 0.0001$, as were the features within artifacts.

For an item analysis, a two-way ANOVA was conducted with the type of feature and the type of category as between-item variables. Again, there was a reliable interaction effect, $F(2,24) = 38.68$, MSe = 37.65, $P < 0.001$, as well as a reliable main effect of the type of feature, $F(2,24) = 30.52$, MSe = 29.71, $P < 0.001$. There was no reliable main effect of the type of category.

2.1.7. Correlation between categorization and causal ratings

Of most importance is whether the causal ratings and the categorization ratings correspond. A correlational analysis was conducted to compare the causal ratings with the categorization ratings for each item and each feature type. Because there were 10 categories with three features in each category, 30 causal judgment-categorization rating pairs were used. The correlation between causal ratings and categorization ratings was reliably high, $r = 0.73$, $P < 0.01$. Fig. 2 shows a scatterplot of 30 pairs of judgments. The numbers above the data points in Fig. 2 correspond to the feature numbers in Table 1.

In order to insure that this positive correlation was not due to only one type of category, two additional correlational analyses were carried out, one for natural kinds and another for artifacts with 15 pairs of judgments for each kind. Both correlations were reliably positive (for natural kinds, $r = 0.83$, $P < 0.01$; for artifacts, $r = 0.66$, $P < 0.01$). The difference between these two correlations was not statistically reliable, using Fisher's Z-transformation, $P = 0.16$. Overall, the results showed that the more causal a feature is, the more important the feature is in categorizing an object. There were positive correlations between causal judgments and category judgments across all items, and within natural kinds and artifacts.

2.1.8. Feature selection

Causal scores of features are sensitive to which features are included in the same set. For instance, feature A might cause many features of Category X. But if none of the features that A causes happened to be included in an experiment, feature A would receive a causal score lower than it actually is because causal scores were calculated based only on the features present in the experiment. Further suppose that feature A's centrality is judged to be high. Then, the correlation between the causal score and feature centrality would become lower than it actually is. In this case,

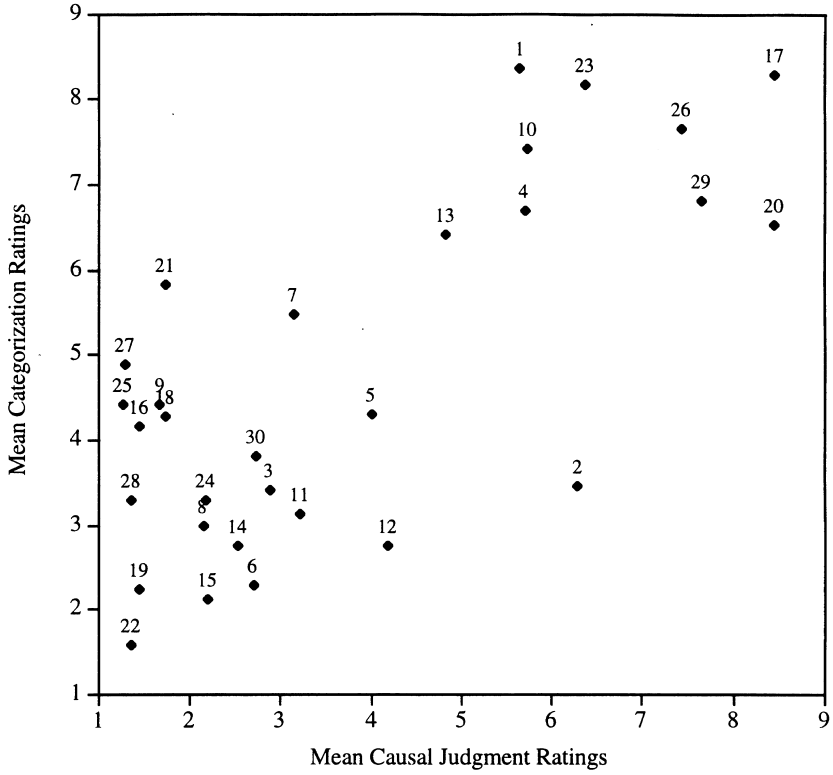

Fig. 2. Relationship between causal judgment and categorization in Experiment 1.

distorted causal scores could lead to a conclusion that the causal status hypothesis is inaccurate when, in fact, it is accurate. Conversely, if feature A's centrality is judged to be low, the correlation between causal scores and feature centrality would become higher (i.e. presenting support for the causal status hypothesis) than is truly the case. It is difficult to judge whether the particular set of features selected by Barton and Komatsu (1989) and used in Experiment 1 was selected particularly in favor of, or against, the causal status hypothesis. Indeed, this potential problem underlying feature selection was the reason why features used in other researchers' studies were used, rather than hand-picked from real-life categories. Because the features were selected by other researchers who were supporting an alternative hypothesis, it seems safe to assume that any errors resulting from feature selection would have been randomly distributed in both directions. Assuming such random sampling, Experiment 1 showed that, overall, the causal status of features accounts for a fair amount of feature centrality in real-life categories. Still, it is important to examine a wide variety of stimulus materials, as will be seen in the next experiment.

2.2. Experiment 2

Counter to the results of Barton and Komatsu (1989) showing that physical appearance matters much less than functional features in artifact categorization, Malt and Johnson (1992) have found that the physical appearance of objects is more important than functional features in artifact categories. At the surface level, these two studies seem contradictory to each other. However, in the framework of the causal status theory, they are not necessarily conflicting because the particular physical features chosen as the stimulus materials in each study could have varied with respect to causal status. Malt and Johnson (1992) used a set of complex physical features, some of which seem causally connected to other features, and some not. For example, a taxi's physical feature was 'has a meter for fare, two seats, and is painted yellow' and its function was 'to provide private land travel for 1–4 people at a time when their own cars are unavailable and they are willing to pay a variable amount of money depending upon their specific destination(s)'. Of the features that were classified as physical ones in Malt and Johnson (1992), 'has a meter for fare' seems to determine the taxi's function whereas 'painted yellow' does not. Based on this observation, the causal status hypothesis can explain the apparent discrepancy in the following way. The physical features used in Barton and Komatsu (1989) were not causal (as has already been shown in Experiment 1), whereas at least some of the physical features used in Malt and Johnson (1992) appear to be more causal.

In order to test this hypothesis, Experiment 2 examined physical features used in Malt and Johnson (1992). Participants received all features used in Malt and Johnson (1992) and indicated causal relations among the features within each category. As in Experiment 1, the centrality of features in categorization was also measured. It was predicted that not all physical features would be equally central, and furthermore, that the centrality of physical features would correlate with their causal centrality.

Hence, the purpose of Experiment 2 is not a comparison between different types of features (e.g. functional versus physical features), but rather to examine the variance within the same type of feature, namely, physical features in artifacts.

2.2.1. Participants

Seventeen undergraduate students at Yale University participated in this experiment. Twelve participated in partial fulfillment of the requirements for an Introduction to Psychology course and five received $7.00 for participating in this experiment and other unrelated experiments.

2.2.2. Materials

The stimulus materials used in Experiment 2 are presented in Appendix A. These features were taken from Malt and Johnson (1992). In their study, a set of multiple physical features was presented as a single item for each artifact (e.g. For desk, 'this thing has a flat, rectangular surface with drawers underneath, four legs, and a matching chair'). These physical features were broken down into separate features (e.g. For desk, 'the surface is flat', 'the surface is rectangular', 'has drawers underneath', 'has four legs', and 'has a matching chair'), because Experiment 2 investigates how these physical features for the same object might vary in centrality. Functional features were also divided if they were connected by the word 'or' in Malt and Johnson (1992; e.g. for boots, 'to protect the feet from the elements, or to prevent excessive strain on the feet, especially while hiking' was broken down into two features). This gave a total of 64 features for 12 artifacts, with 4–7 features for each artifact. There were 49 physical features and 15 functional features.

A set of 49 categorization questions was developed for the physical features. Each question followed the format 'would X still be X if it were in all ways like X except that it does not have property Y?' where X was a category, and Y was one of its physical features. There were no categorization questions for functional features because Experiment 2 examined physical features alone rather than making a comparison between the two types of features. The order of categorization questions was randomized across all participants. The task was programmed using MacProbe (Hunt, 1994).

In measuring the causal status of features, a task different from Experiment 1 was used because the total number of all possible pairwise bi-directional causal questions is 284, too many for participants to complete in a reasonable amount of time. Instead, the following task, adapted from Sloman et al. (1998) and Sloman and Ahn (1998), was used. An object's features were arranged on a sheet of paper in a circle. The position of features was randomized across all objects and across all participants. At the top of each sheet was the name of the artifact. All participants received all 12 artifacts in a randomized order. The instructions were as follows. (The instructions used an example from a natural kind in order to avoid any confounds that might result from using an artifact example.)

> Objects in the world are comprised of many features. For instance, some features of the object 'BIRD' are that it 'has wings', 'flies', 'has bird DNA',

'chirps', and so on. Some of the features of 'BIRD' determine other features. For example, the feature 'has wings' permits or allows birds to 'fly'. In this experiment, you will be given the name of an object and a number of features for that object. If you take a look at the green booklet placed next to your computer, an object's name is placed on the top of each page and the features are arranged in a circle.... Within each object, please determine which features cause, determine, permit, or allow which other features by drawing an arrow between the features. For example, if you believe 'having wings' allows birds to 'fly', please draw an arrow as follows.

have wings ——————— → fly

You may draw any number of arrows from any number of features to any number of features. That is, one feature may cause or determine two or three features and one feature may be caused or determined by two or three features. In addition to drawing lines and arrows to designate causal relations, you should also put a number by each arrow you draw in order to express how strong or weak you think that relationship is. You should use a scale from 1 to 5.

1	2	3	4	5
Very weak	Weak	Neither weak nor strong	Strong	Very strong

Please rate the strength of ALL of the arrows you draw. What if you think there are other important features that have not been included? You are free to add these yourself. Just write the name of the feature in an open space and draw arrows to or from that feature as usual, and remember to include a numerical rating for the relationship.

In order to obtain causal structures that participants spontaneously draw, the instructions further emphasized that it is acceptable not to draw any causal relations if he/she believed that there was no causal relations in the object. In that case, participants were asked to write down 'none' on the sheet so that it could be distinguished from a missing response.

2.2.3. Procedure

The presentation order of the causal judgment task and the categorization task was counterbalanced. Participants were randomly assigned to one of the two presentation orders. Eight participants performed the categorization task first, and nine performed the causal judgment task first. Participants were first seated in front of a computer, which displayed the instructions for the task they were assigned to perform first. For the causal judgment task, the instructions were displayed on a computer screen and the participants performed the task on a booklet. For the categorization judgment task, all questions were displayed on a computer in a randomized order.

2.3. Results and discussion

The pattern of the results remained the same regardless of the order in which participants performed the task. Hence, all results were collapsed in the following analyses. In all of the analyses and figures reported below, the ratings on the categorization task were inverted as in Experiment 1 for ease of understanding, so that higher ratings indicate higher centrality.

On average, a participant drew 3.8 causal links per object. Each physical feature's causal status was calculated by obtaining the total strength of links in which the feature serves as a direct cause within that category, divided by the number of features in the category minus 1 (i.e., the total number of features a feature can cause within a category). For instance, suppose A causes B with a strength of 3, and A causes C with a strength of 2. Suppose that there were four features for this object (A, B, C, and D). Then, A's causal score was (3 + 2)/3, and the causal scores of B, C, and D were all 0. A correlational analysis was carried out for 49 pairs of causal scores and ratings on the categorization task for physical features, resulting in a reliably positive correlation of 0.53, $P < 0.001$. Another correlational analysis was conducted between the summed causal scores (i.e. the total causal strengths of all links in which a feature participates as a direct cause) and the ratings on the categorization task, and showed a reliably positive correlation of 0.60, $P < 0.001$. Hence, Experiment 2 found that even within the same type of feature, the causal status of features correlated with their conceptual centrality.

Fig. 3 shows a scatterplot of 49 pairs of judgments. The causal scores plotted in

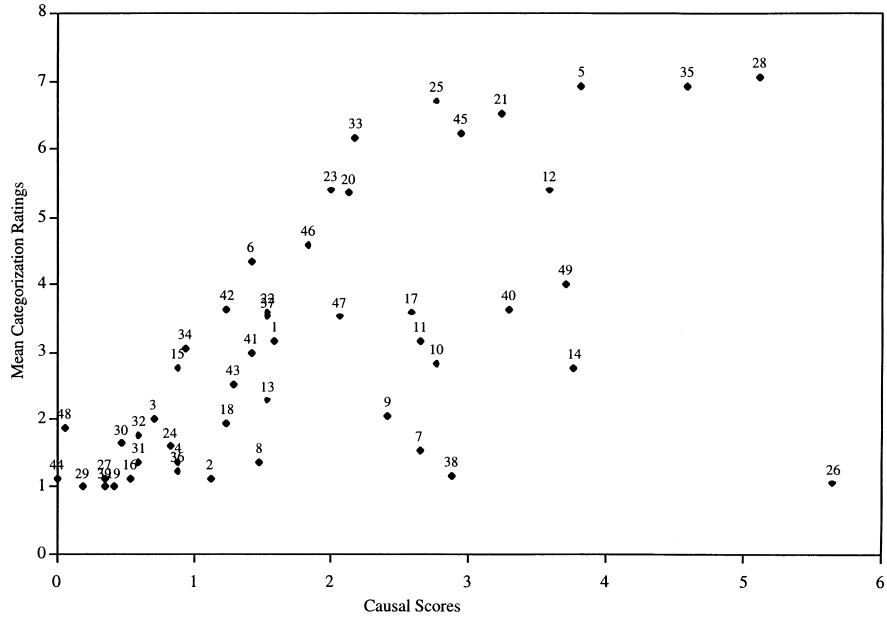

Fig. 3. Relationship between categorization and causal judgment in experiment 2.

Fig. 3 are the sum causal scores. The numbers above the data points in Fig. 3 correspond to the feature numbers in Appendix A. As can be seen in Fig. 3, one obvious outlier is Feature 26, which is 'is 12 inches long' for a ruler. This feature was judged to be high in causal status but low in concept\ual centrality. This outlier seems to have occurred because this feature also had low category validity, in that the length of rulers vary a lot. Hence, this would be an example of the case where a counteracting force (e.g. low category validity) can override the causal status effect. When this outlier was eliminated from the analysis, the correlation between the causal and categorization ratings was increased to 0.68 based on the average causal scores, and 0.74 based on the summed causal scores.

3. Part B: direct manipulation of causal status

So far Experiments 1 and 2 have shown strong correlations between the causal status of features and feature weighting in natural kinds and artifacts. However, these experiments have not yet established the causal connections between these two factors. For instance, the high correlations can mean that the causal status of features directly determine feature weighting as the causal status hypothesis states. Unfortunately, they can also mean that more important features are perceived to be more causal. That is, people might treat molecular features in natural kinds and functional features in artifacts as more important features as the domain-specific hypothesis states, and then because they are central features, they might be perceived to be more causal. As a result, we have not yet nailed down the reason why the interaction between the category type and the feature type has been obtained. Although there are merits in examining real-life categories, the only way to directly test the causal status hypothesis is to use artificial stimuli and experimentally manipulate the causal status of features in novel categories.

The two experiments presented in Part B used artificial categories in order to pit the causal status of features (whether a feature was a cause or an effect of another feature) against type of category (natural kind or artifact) and type of feature (compositional or functional). First, novel objects are created to be either natural kinds or artifacts. Second, each object has two types of features, those which showed the category-feature interaction effect in previous studies. One type is functional, or what an object does (e.g. a mirror can reflect an image) or is used for (e.g. a chair is used for sitting). The other type is a broader class which includes molecular structure (e.g. has goat DNA), internal parts (e.g. has a spleen inside), substance (e.g. is made of wood), and origin (e.g. was born from cow parents). All these features have been shown to be essential in natural kinds, although different researchers have focused on different types of features. For example, Barton and Komatsu (1989) focused only on molecular features whereas Gelman (1988) and Keil (1989) have also examined internal parts (e.g. has spleen inside) or origins (e.g. a guitar used to be a larch). To refer to all of these features, the term 'compositional features' will be used from now on. In the stimulus materials used in Part B, each object, either a natural kind or an artifact, possesses two features, one compositional

and one functional. The third manipulation of Experiments 1 and 2 is the causal status of features. In one condition the compositional feature causes the functional feature, and in the other, the compositional feature is caused by the functional feature.

The prediction is that when a compositional feature is caused by a functional feature in both natural kind and artifact categories, the functional feature will be more important for categorization than the compositional feature. This result would contrast some previous findings for natural kinds, and furthermore, would present direct evidence against the hypothesis stating that compositional features are always essential features for natural kinds. Similarly, when a functional feature is caused by a compositional feature in both types of category, it is predicted that the compositional feature will be more important for categorization than the functional feature.

Several special measures were taken in both experiments to eliminate potential confounding variables. To minimize item differences between cause and effect features, the same functional feature served as either a cause or an effect feature of the same molecular feature. For instance, in one condition, Kehoe ants had blood high in iron sulfate (compositional feature) which caused fast food digestion (functional feature). In the other condition, fast food digestion in Kehoe ants caused blood high in iron sulfate. This measure was taken to insure that the causal status effect does not occur for reasons other than the causal status of the features per se.

Experiments 3 and 4 also controlled for two other possible determinants of feature centrality. The category validity (i.e. the probability of having a certain feature given that an object belongs to a certain category) was held constant by explicitly giving fixed probabilities at the beginning of each feature description. The cue validity (i.e. the probability of belonging to a certain category given that an object has a certain feature) of cause features was indirectly equated with the cue validity of effect features by using identical features for cause and effect features across two versions. That is, in one version feature A caused feature B, and in the other version feature B caused feature A. When data from the two versions were collapsed, the cue validity of the cause feature was the cue validity of features A and B. Likewise, the cue validity of the effect feature was the cue validity of features A and B, resulting in the same cue validities for the cause and effect features. In this way, any results from Experiments 3 and 4 cannot be due to cue or category validities of features.

3.1. Experiment 3

In addition to the controls described in the previous section, Experiment 3 equated features used for artifacts and natural kinds as much as possible. For instance, in the natural kind condition a Kehoe was an ant with blood high in iron sulfate and the capacity to digest food quickly, whereas in the artifact condition, a Kehoe was a septic system with liquid high in iron sulfate and the capacity to decompose sewage quickly. (See Appendix B for a complete list of features.) Although some changes in features were unavoidable as in this example, the item differences across the two domains were minimized as much as possible. This is an important measure to take in attempting to demonstrate the effect of different beliefs underlying artifacts and

natural kinds (e.g., a belief that different types of feature are essential for different domains). No previous studies showing the domain effect have adopted such procedures, making it difficult to interpret their results because the effect could have been due to systematic variations accompanying item differences.

3.1.1. Participants

Twenty-eight undergraduate students at Yale University participated in this experiment in partial fulfillment of requirements for an Introduction to Psychology course.

3.1.2. Materials and design

Four novel objects were created, and named Oenotheras, Kehoe, Yabuka, and Coryanthes. Each passage describing an object began with a sentence stating whether the object was a natural kind or an artifact. (See Appendix B for the materials used in Experiment 3.) For instance, in the natural kind version of Kehoe, they were told, 'there is a species of ant called Kehoe ants'. For an artifact version, they were told, 'there is a kind of septic system called Kehoe septic'. In developing natural kind and artifact versions of the stimuli, various types of natural kinds and artifacts were used to assure generality of the results. For natural kinds, there were two plants, one animal, and one mineral. For artifacts, there was one moving machine, one stable system, one synthesized material, and one type of house.

Followed by the statement specifying natural kind/artifact status were two sentences describing the object's compositional feature (e.g. 'have blood high in iron sulfate' for Kehoe ants) and functional feature (e.g. 'digests food fast' for Kehoe ants). In Experiment 3, all compositional features were molecules. The category validity of each feature was described to be 80%. For instance, they were told that '80% of Kehoe ants have blood high in iron sulfate. 80% of Kehoe ants digest food fast'. The order in which the compositional and functional features were presented within the same object was counterbalanced across all manipulations of the experiment.

After these two sentences about the category validity information of two features, participants were presented with a paragraph on how these features were causally related. There were two conditions, depending on which feature serves as a cause. In the F-Cause condition, the functional features were described as causing the compositional features. In the C-Cause condition, the same compositional features were described as causing the functional features.

Each paragraph describing the causal relations was constructed using the following format. First, there was a statement about the causal relation (e.g. 'Kehoe ants' fast food digestion tends to cause blood high in iron sulfate'). Then, there were one or two sentences explaining the causal mechanism (e.g. in the F-Cause condition, 'Fast food digestion increases Kehoe ants' metabolism which results in blood high in iron sulfate', and in the C-Cause condition, 'The iron sulfate in Kehoe ants' blood stimulates the enzymes responsible for manufacturing the food-digesting secretions, and a Kehoe ant can digest food faster with more secretions'). In developing this

second part, cause features were mentioned as frequently as effect features in order to equate discourse saliency. In addition, special care was taken to equate the natural and artificial kind versions as much as possible, while still making the description as plausible as possible. Two undergraduate students who were unaware of the hypotheses of the experiments judged the plausibility of these two versions from all four objects, and found no differences among these eight versions with respect to plausibility (mean of 5.3 on the materials in the C-Cause condition and 5.4 on the materials in the F-Cause condition on a 1 (implausible)–7 (plausible) rating scale).

Following each passage, there were two questions about the likelihood that some object was a member of the novel category described in the passage. The object in one question possessed the compositional feature, but not the functional feature, of the novel category. The object in the other question possessed the functional feature, but not the compositional feature, of the novel category. In this way, each question measured the centrality of each feature. The questions following each passage are also presented in Appendix B. The order in which the two questions were presented to each participant was counterbalanced across all conditions.

In sum, 64 passages were developed by crossing 4 objects (Oenotheras, Kehoe, Yabuka, and Coryanthes), two Category Types (natural kinds and artifacts), two Causal Directions (C-Cause and F-Cause), two orders of feature descriptions (C-first and F-first), and two orders of questions (C-first and F-first). Then, for each participant, a booklet containing four passages was prepared with the following restrictions. All four passages were about different objects; two passages were F-Cause and two were C-cause; two passages were F-first and two were C-first with respect to feature presentations; two were F-first and two were C-first with respect to the question presentations; and finally, two were natural kinds and two were artifacts. This way, participants received all possible manipulations but never saw the same object more than once. The order of the four passages within each booklet was randomized.

The critical variables were Category Type (natural kinds and artifacts), Causal Direction (C-Cause and F-Cause), and Question Feature Type (C or F). Therefore, Experiment 3 was a $2 \times 2 \times 2$ within-subject design.

3.1.3. Procedure

Each participant received a booklet with four short passages and two questions following each passage, as was described in the Materials section. Participants were instructed to read each passage, and to answer each question by writing down a number from 0 to 100 to represent the likelihood that the object described in the question was a member of the novel category described in the passage. They were further instructed that 100 would mean that the object was definitely an example of the novel object, 0 would mean that it definitely was not, and 50 would mean that it was half and half.

3.1.4. Results

Average likelihood ratings from natural kinds and artifacts are shown in Fig. 4, along with the standard errors as represented by the error bars. For both natural kind

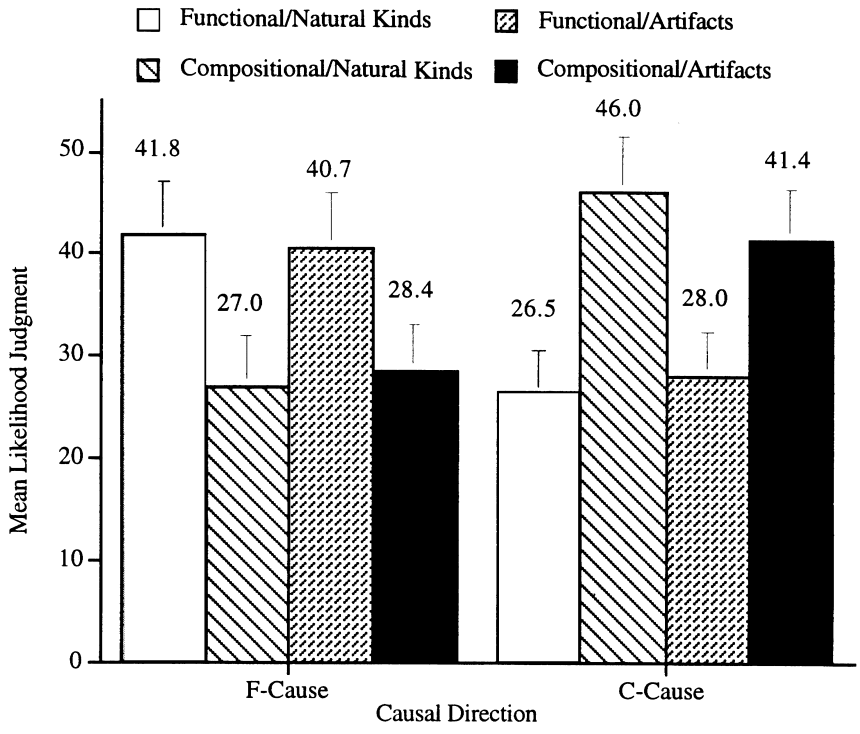

Fig. 4. Mean ratings on natural kinds in Experiment 3.

and artifact categories, causal status, rather than feature type, was used as a basis for categorization. Regardless of the type of category, the F-Cause condition led to higher ratings for objects possessing the functional feature than objects possessing the compositional feature (14.8% more for natural kinds and 12.3% more for artifacts). The C-Cause condition led to higher ratings for objects possessing the compositional feature than for objects possessing the functional feature (19.5% more for natural kinds and 13.4% more for artifacts). Collapsing across the F-Cause and C-Cause conditions, ratings on objects possessing a compositional feature were essentially the same as ratings on objects possessing a functional feature, both for natural kinds (36.5% and 34.2%, respectively) and for artifacts (34.4% and 34.9%, respectively)[2].

A three-way ANOVA was conducted on likelihood ratings with Category Type

[2]One might be concerned that overall mean ratings are below 50%. Indeed, these means seem quite reasonable considering that each category was described to have only two characteristic features and the category validity of each characteristic feature was described to be 80%. Hence, the absence of one feature is expected to bring down the mean membership likelihood to below 50%. In addition, it should be noted that the causal statue hypothesis does not claim that a cause feature is a defining one, and hence the fact that the ratings for items with a cause feature were below 50% does not pose a problem for the causal status hypothesis. (See the general discussion for more detail.)

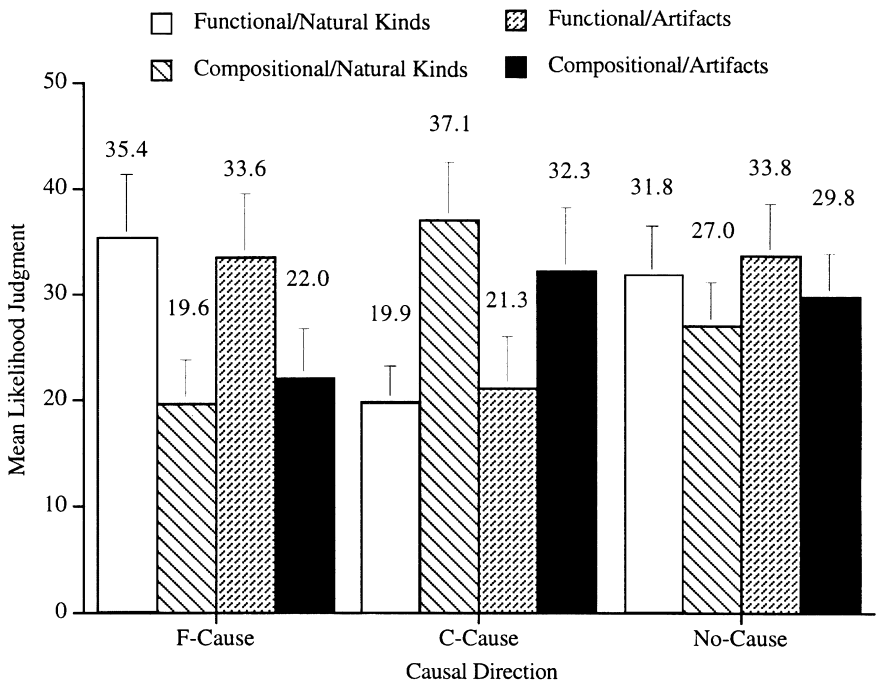

Fig. 5. Mean ratings on artifacts in Experiment 4.

(natural kind or artifact), Causal Direction (F-cause or C-cause), and Question Feature Type (compositional or functional) as within-subject variables. As predicted, the only effect shown to be reliable was the interaction effect between Causal Direction and Question Feature Type, $F(1,27) = 16.93$, MSe = 746.1, $P < 0.001$, such that both compositional and functional features were more important for categorization when they were causes than when they were effects. There were no other reliable effects, $P > 0.40$. Most importantly, there was no reliable interaction effect between Category and Question Feature Type, $F(1,27) = 0.09$, $P = 0.77$. The lack of an interaction effect between Category and Question Feature Type suggests that previous results showing a compositional feature bias for natural kinds and a functional feature bias for artifacts were likely to have been driven by the causal status of the features rather than by the type of feature per se.

3.1.5. Item analysis

An item analysis was carried out for four natural kinds and four artifacts. A three-way ANOVA was conducted with Category type (natural kinds and artifacts) as a between-item factor, and Causal Direction and Question Feature Type (functional and compositional) as within-item factors. Again, the only reliable effect was the interaction between Causal Direction and Question Feature Type, $F(1,6) = 29.67$, MSe = 61.3, $P < 0.01$. None of the main or interaction effects were reliable, $P > 0.50$.

3.1.6. Order effect

One might be concerned that the participants may have noticed the manipulation of the causal relations among features across the four passages, and therefore may have showed the causal status effect due to demand characteristics. In order to rule out this alternative account, the ratings on the passage presented as the first item for each participant were examined. The pattern was consistent with the overall results. In the F-Cause condition, the mean rating on the item with only the functional feature (38.9%, $n = 12$) was higher than the mean rating on the item with only the compositional feature (23.8%). In the C-Cause condition, it was the opposite; the mean rating on the item with only the compositional feature (45.6%, $n = 16$) was higher than the mean rating on the item with only the functional feature (31.1%).

3.2. Experiment 4

Experiment 4 tests the same hypothesis as in Experiment 3, using different materials. Experiment 3 provided the strongest possible test of the hypothesis by using the same features for both natural kinds and artifacts and, furthermore, by using the same features for both the F-Cause and C-Cause conditions. Unfortunately, it is extremely difficult to develop many sets of stimulus materials that could be used for both natural kinds and artifacts, and are at the same time reversible across the two causal direction conditions. For that reason, only four items were developed for Experiment 3. Furthermore, the materials used in Experiment 3 may have been restricted by the constraint to use the same features for both natural kinds and artifacts. In order to obtain more generality of the findings across many sets of stimulus materials, Experiment 4 used materials and categories different from Experiment 3, and used different features for natural kinds and artifacts. In addition, Experiment 4 included a condition in which no causal relations were specified (No-cause condition). The results from this condition could determine whether the two features used in each example are equally salient or central when no causal status is given to the features. Other than these changes, the design and the hypothesis of Experiment 4 were identical to those of Experiment 3. It was predicted that the centrality of features would be determined as a function of causal status rather than as a function of whether the feature is compositional or functional, in natural kinds or in artifacts.

3.2.1. Participants

Twenty-seven students at Yale University participated in this experiment in partial fulfillment of requirements for an Introductory Psychology course.

3.2.2. Design and materials

Passages describing six novel objects were developed. Three of the objects were natural kinds (Natural Kinds 1–3) and three were artifacts (Artifacts 1–3). (See Appendix C for a complete list of all six objects, their features, and the questions asked about each object.) As in Experiment 3, these six objects were selected to cover a wide variety of natural kinds and artifacts. The categories used were animal,

fruit, and flower for natural kinds and computer, artificial flavor, and machine for artifacts.

Each object had two features, a compositional feature and a functional feature. Because one of the purposes of Experiment 4 was to obtain generality of the results of Experiment 3 over stimulus materials, none of the features used in Experiment 3 was used in Experiment 4. Different features were used for each of the six objects. The two features of each object were selected such that they would be equally salient when the features were not causally connected to each other, as determined by results of a norming study[3]. Along with a description of the two features possessed by the object, each passage contained a statement about the category validity of each feature as in Experiment 3 (i.e. 80% of the members in the target category possess the compositional feature; 80% of the members in the target category possess the functional feature). The order in which the compositional and the functional feature were mentioned in the passage was counterbalanced.

For each object, three passages were developed, depending on the way the compositional and functional features were related: F-cause (the functional feature caused the compositional feature), C-cause (the compositional feature caused the functional feature), and No-cause (there was no causal relationship between the two features). As in Experiment 3, the same features were used for the different Causal Direction conditions so that the only difference between the cause and the effect features would be their causal status. For the No-cause passages, there was no additional paragraph following the feature descriptions. For the F-cause and C-cause conditions, one paragraph consisting of two or three sentences was added after the feature descriptions. (see under F-cause and C-cause for each object for verbatim paragraphs.) In developing these paragraphs, special care was taken not to emphasize any single feature by mentioning all features with equal frequency. As in Experiment 3, each paragraph started out with a statement about the causal relation between two features of an object followed by a brief account of the mechanisms.

Each participant was randomly assigned to one of three Passage Sets (1–3), each of which contained six passages. Within each set, there were two passages containing each of the three relations (C-cause, F-cause and No-cause), one passage describing a natural kind and one describing an artifact. Each relation occurred in different passages across the three Passage Sets. Set 1, for example, had Natural Kind 1 in the F-cause form, Natural Kind 2 in the C-cause form, Natural Kind 3 in the No-cause form, Artifact 1 in the F-cause form, Artifact 2 in the C-cause form, and Artifact 3 in the No-cause form. This way, all participants saw one passage from each of the six conditions (three causal background $\times$ two types of category) but never saw the same object more than once. A Latin square design was used to determine which natural kinds and artifacts should have which relation in each set.

Following each passage, there were two questions about the likelihood that some object was a member of the novel category described in the passage, as in Experi-

[3]The norming study was informally conducted by asking 2 undergraduate students which feature seemed more important to them in categorizing the target objects. See the results from the No-Cause condition for further support.

ment 3. The questions following each passage are also presented in . The order of the two questions was counterbalanced across passages.

In sum, the design of Experiment 4 was a $3 \times 2 \times 2$ within-subjects design. The three variables were Causal Directions (C-cause, F-cause, and No-cause), Category Type (natural kind and artifact), and Feature Type (functional and compositional).

3.2.3. Procedure

The procedure was identical to that used in Experiment 3.

3.2.4. Results and discussion

The results replicated those of Experiment 3. The causal status of features, rather than type of feature or type of category, determined feature importance for categorization. Fig. 5 shows the mean ratings from natural kinds and artifacts, respectively, along with the standard errors indicated by error bars. In the F-Cause condition, objects possessing a functional feature received higher likelihood ratings than those with a compositional feature (15.8% higher for natural kinds and 11.6% higher for artifacts). In the C-Cause condition, objects possessing a compositional feature received higher likelihood ratings (17.2% higher for natural kinds and 11.0% higher for artifacts). Collapsing across the F-Cause and C-Cause conditions, ratings on objects possessing a compositional feature were essentially the same as ratings on objects possessing a functional feature, both for natural kinds (28.5% and 27.7%, respectively) and for artifacts (27.2% and 27.5%, respectively).

A three-way ANOVA was conducted with Category Type (natural kind or artifact), Background Knowledge (F-cause, C-cause, or No-cause), and Question Feature Type (compositional or functional) as within-subject variables. As predicted, the interaction between Causal Direction and Question Feature Type was statistically reliable, $F(2,52) = 12.41$, MSe = 436.58, $P < 0.001$. That is, feature importance varied depending on causal status. No other effect was reliable, $P > 0.50$. Most importantly, there was no reliable interaction effect between Question Feature Type (compositional or functional) and Category Type (natural kinds and artifacts), $F(1,26) = 0.01$, $P = 0.92$. In addition, the centrality of compositional and functional features was compared in the No-Cause Condition to examine whether there was any a priori feature centrality. There was no statistically reliable difference between the two types of features for natural kinds, $t(26) = 0.98$, $P = 0.34$, or for artifacts, $t(26) = 0.86$, $P = 0.40$.

3.2.5. Item analysis

An item analysis was carried out for three natural kinds and three artifacts. A three-way ANOVA was conducted with Category Type (natural kinds and artifacts) as a between-item factor, and Causal Direction and Question Feature Type (functional and compositional) as within-item factors. Again, the only reliable effect was the interaction between Causal Direction and Question Feature Type, $F(2,8) = 23.21$, MSe = 26.6, $P < 0.001$. None of the main or interaction effects were reliable, $P > 0.40$. Again, there was no statistically reliable difference between the functional and the compositional features in the No-Cause condition, $t(5) = 1.16$,

$P = 0.30$. These results strongly support that features equal in centrality when causal relations were not given can show varying centrality when causal relations were given.

3.2.6. Order effect

As in Experiment 3, the responses made on the first passage were examined to see whether their initial responses follow the overall pattern, and to ensure that the overall results were not due to demand characteristics resulting from repeated manipulations of the same variable. The mean ratings on the first passage were consistent with the overall pattern. In the F-Cause condition, the mean rating on the item possessing the functional feature (42.6%, $n = 8$) was higher than the mean rating on the item possessing the compositional feature (31.7%). In the C-Cause condition, the opposite occurred; the mean rating on the item possessing the compositional feature (35.0%, $n = 11$) was higher than the mean rating on the item possessing the functional feature (24.5%). In the No-Cause condition, the mean rating on the functional features (22.0%, $n = 8$) was the same as the mean rating on the compositional features (22.6%).

In summary, Experiment 4 replicated Experiment 3. The causal status effect was obtained in both studies suggesting that the interaction between the category and feature type found in previous research (e.g. Barton and Komatsu, 1989) may be due to the causal status of the features, rather than the type of category or feature per se, and that there is nothing special about compositional features of natural kinds or functional features of artifacts when causal status is equated.

4. General discussion

That people weight causes more than effects provides one account of why people treat artifacts and natural kinds differently in certain tasks. Part A examined the extent to which the causal status hypothesis could account for the variance in real-life categories. Experiments 1 and 2 measured people's existing causal knowledge of categories that were shown to exhibit differential weightings of features in previous research. As predicted by the causal status hypothesis, features seen as more important were also judged to be the causes of features seen as less important. Such findings were obtained with both natural kinds and artifacts consisting of functional, physical, and molecular features (Experiment 1). The causal status hypothesis further accounts for the apparent conflicting results of Barton and Komatsu (1989) and of Malt and Johnson (1992). Whereas Barton and Komatsu (1989) found that functional features are more essential than physical appearance, Malt and Johnson (1992) found that physical features are more important than functional features in artifacts. The physical features used in Barton and Komatsu (1989) were not causally central (Experiment 1) whereas some of the physical features used in Malt and Johnson (1992) seemed to play causal roles (Experiment 2). Experiment 2 indeed found that even among physical features, the causal status of features reliably predicted feature centrality.

The two sets of experiments reported in Parts A and B complement each other. Experiments 1 and 2 showed strong positive correlations between the causal status of features and feature weighting in real-life categories. However, the use of real categories did not allow perfect control over features. There are other determinants of features weights, such as category validity and cue validity, that could not be controlled for in the real-category experiments. Furthermore, correlational studies cannot demonstrate whether causal status per se can induce feature centrality. Experiments 3 and 4 used artificial stimuli to control for cue and category validities, directly manipulated the causal status of novel features, and demonstrated that causal status indeed can determine feature centrality.

The following section will address three issues. First, the causal potency of features has been frequently mentioned in the context of essentialism. The causal status hypothesis will be compared with essentialism, and the specific points of agreement and disagreement of these two views will be described. Second, the current study examined only one difference between natural kinds and artifacts. The implications of the current results on other differences between natural kinds and artifacts will be discussed. Third, implications for developmental studies and models of categorization will be presented.

4.1. Essentialism

One traditional and pervasive view about concepts is that things have essences that are deeper and more basic to a kind. According to Locke (1894/1975), 'Essence may be taken for the very being of any thing, whereby it is, what it is (p. 417).' Essentialism is particularly related to the causal status hypothesis in that they both concern the causal potency of essential (or central) features. Locke continues, 'And thus the real internal, but generally in Substances, unknown Constitution of Things, whereon their discoverable Qualities depend, may be called their Essence' (p. 417). Similarly, Putnam (1977) states, 'If I describe something as a lemon, or as an acid, I indicate that it is likely to have certain characteristics (yellow peel, or sour taste in dilute water solution, as the case may be); but I also indicate that the presence of those characteristics, if they are present, is likely to be *accounted for* by some 'essential nature' which the thing shares with other members of the natural kind (Putnam, 1977, p. 104, emphasis added)'. Medin and Ortony (1989) also proposed psychological essentialism: Whether or not things actually have essences, people act as if things have essences which are responsible for surface features of objects. Atran (1987) suggested that in biological kinds this belief is universal, and Gelman and Wellman (1991) further demonstrated that this belief in essences is present even in young children. In sum, essentialism argues that there is an essential nature which accounts for or determines what the entity's superficial properties will be. That is, both essentialism and the causal status hypothesis propose that essential (or central) features have causal potency.

At the same time, essentialism differs from the causal status hypothesis along

[4]It should be emphasized that the causal status hypothesis departs from specific versions of essentialism (as noted by the references below) rather than essentialism as a whole.

several important dimensions[4]. (1) The causal status hypothesis does not necessarily assume that causal features are defining features. (2) The causal status hypothesis does not dichotomize features into essential and surface features. (3) The causal status effect arises as a result of specific knowledge people have about causal relations, whereas some essentialists argue that essential properties are independent of our knowledge of them. (4) The causal status hypothesis is expected to operate with both natural kinds and artifacts whereas some essentialists argue that artifacts do not have real essences. Each of these points will now be discussed in more detail.

4.1.1. Essences as defining features for concepts

As pointed out above, an object's essence is believed to be the very being of what it is, and hence, essences are generally considered to be defining features of a category (see Keil, 1989, p. 37 for a similar interpretation with respect to Locke's nominal essences; also the strong version of psychological essentialism described in Malt, 1994). According to this version, if an object has an essence of a certain kind, the object must be a member of the kind, and if an object is a member of a certain kind, the object must possess the essence: essences are necessary and sufficient features of a kind.

In contrast, the causal status hypothesis does not make any statement about whether or not there is a certain type of features that serves as defining features. That is, the causal status hypothesis should not be taken as a claim that a cause feature is a defining feature of a category. Indeed, such a claim would not make any sense because a feature's causal status, by definition, is relative in that a feature which is a cause of one feature can be an effect of another feature. For instance, wings in birds are the effect of bird DNA, and at the same time, wings allow birds to fly. Therefore, features cannot be simply partitioned into cause features and effect features, and for that reason, it would be absurd to claim that cause features are defining features. Although it might be possible to conjecture that the most terminal cause (which is an essence in the essentialist framework) serves as a defining feature, the causal status hypothesis in its current form is mute about the debate on whether or not concepts have (or are believed to have) defining features.

4.1.2. Dichotomy between essential features and surface features

While some essentialists explicitly note that feature centrality should be thought of as a continuum (e.g. Medin and Ortony, 1989), other essentialists (e.g. Putnam, 1977) tend to dichotomize features into essential ones and surface ones. That is, the emphasis has been on the difference between deep features and surface features. As a result, essentialism fails to offer an account of how surface features might vary in centrality. In contrast, the gradient structure of feature centrality is built into the causal status hypothesis because the hypothesis is about the relative difference between causes and their effect feature. As described in the introduction to this paper, Ahn et al. (in preparation) provide empirical evidence for this: When A causes B, and B causes C, the centrality of these features were ranked in the order of A, B, and C.

In addition, essences are often described as internal, hidden, or unobservable

properties such as atomic weight, genetic structure, and so on, whereas surface features are external or perceptual features such as wings and feathers in birds (e.g. Putnam, 1975). Locke (1894/1975) even stated that real essences are undiscoverable. It is not clear whether these are statements about the way things happen to be in the world (i.e., essences happen to be internal and non-perceptual features, and surface features happen to be perceptual), or statements about what essential or surface features really are. The causal status hypothesis does not impose any restrictions on the type of cause or effect features, or the type of central or peripheral features. That is, a perceptual feature can be judged to be more central than a non-perceptual feature if the former causes the latter. Ahn et al. (in preparation) provide empirical evidence that even when features are all observable symptoms, their causal status determines the relative difference in centrality among them.

4.1.3. Essentialism and specific knowledge

Some essentialists (e.g. Kripke, 1971; Putnam, 1975) have proposed that categorization of natural kinds is independent of our knowledge of essential properties. For instance, even if it turns out that water does not consist of H_2O, and instead, consists of XYZ, what we have been referring to as water is still water. Similarly, even if we discover that all cats are, in fact, robots controlled from Mars, according to the essentialist view, the word cat always refers to the same category of objects. In that sense, natural kinds are like proper nouns in that even if we find out, for example, that Shakespeare did not write Romeo and Juliet, Shakespeare is nonetheless Shakespeare. In contrast, Braisby et al. (1996) demonstrated that after discovering changes of essential properties (e.g. cats being robots controlled from Mars), people responded that category membership of the object should also be somewhat changed (e.g. the object formerly known as a cat is not a cat any more). As discussed earlier, the causal status hypothesis is rather agnostic about the exact characteristics of essences of things, and therefore, the current experiments do not provide any direct evidence for or against Kripke's and Putnam's claims about natural kinds. However, one aspect of the causal status hypothesis that contrasts with Putnam's and Kripke's essentialism is that the causal status hypothesis assumes more responsibility of specific knowledge: Feature centrality would change as our knowledge about the causal relations among features changes, a finding demonstrated in Experiments 3 and 4.

4.1.4. Domain-specific essences: objectivity and the relevance of science

Although Putnam (1975) argues that all kinds (both natural and nominal kinds) have essences, Schwartz (1978) claims that for nominal kinds (including artifacts), there is no real essence. Nominal kinds, such as 'white things', are conventionally established, and therefore, if the criterial properties change (e.g. if an object is stained with mud), it no longer belongs to the category. Schwartz (1979) also argues that there can be real sciences for natural kinds (because the aim would be to discover the real essence of the natural kinds) whereas there cannot be any real science for artifacts because there is nothing to discover in conventionally established artifacts. In general, natural kinds are thought of as being discovered from the world,

whereas artifacts are thought of as being constructed (see Kalish, 1998 for empirical demonstration with children and adults). Although this objectivity might be an important distinction between natural and artifactual kinds, the current study showed that the causal status effect occurs regardless of whether concepts are artifactual kinds or natural kinds. Hence, this distinction between categories with nominal and real essences does not seem to have any impact on the causal status effect.

4.2. Summary

This section provided theoretical clarification of the causal status hypothesis by contrasting it with essentialism. For the purpose of clarification, the discussion has highlighted the differences between these approaches. However, these two views agree on the most important claim: that features that cause other features are more essential or central. Experiments 1 and 2 of the current paper showed that causal centrality is correlated with conceptual centrality in real-life categories. Experiments 3 and 4 showed support for a stronger version of this claim by demonstrating that imposing causal potency on a feature actually causes an increase in its centrality. In this way, the current work can be construed as support for the primary claim of essentialism. Still, other discrepancies between essentialism and the causal status hypothesis reviewed in this section remain to be more systematically examined in the future.

4.2.1. Implications for natural kinds/artifacts distinctions

This section examines the implications of the causal status hypothesis on the distinctions between natural kinds and artifacts. Clearly, the causal status hypothesis and the current results do not eliminate the possibility that there might be other differences between natural kinds and artifacts. Existing natural kinds and artifacts have numerous differences that were not systematically examined in the current study. This section discusses domain differences or domain-related issues that are frequently mentioned in the literature (see Keil, 1989, Chapter 3 for further discussion).

4.2.2. Internal/external

One recurrent intuition which was the focus of the current paper is that internal features are more important for natural kinds whereas external features are more important for artifacts. The introduction to this paper provides several examples of empirical evidence that are consistent with this observation. One interpretation of such results that is certainly in the air (and explicitly endorsed by Barton and Komatsu) is that an object's feature centrality is determined by virtue of the object belonging to a certain domain. In contrast, the current study showed that an object's feature centrality is rather a function of its causal potency than the object's domain per se. When the causal status of features was held constant, the internal/external distinction did not account for feature centrality (Experiments 3 and 4).

It should be further pointed out that the causal status hypothesis is not contradictory to the work of Keil, or the work of Gelman and her colleagues. More

recently, Gelman et al. (1994) reviewed possible ways in which essences might become domain-specific, and favored the account that 'essentialism is a domain-general assumption, but one that gets invoked differently in different domains, responding to the causal structure of each domain' (p. 358). (See also Keil, 1989 for similar discussion on differences in causal structure of natural kinds and artifacts.) Note that this is how the causal status hypothesis, a domain-general mechanism, accounts for the domain-dependent differences. While preferring this over other domain-specific accounts, Gelman et al. (1994) admitted at the time that without further evidence the theory remained speculative. The contribution of the current work (especially Experiments 3 and 4) is to demonstrate exactly how a domain-general mechanism of the causal status hypothesis is able to account for the domain differences, and to provide for the first time the empirical evidence for that account.

4.2.3. The need for intention

Another strong intuition about the distinction between natural kinds and artifacts is that artifacts involve a designer's intention (Keil, 1989). Bloom (1996) proposes the strongest version of this claim: 'the extension of artifact kind X to be those entities that have been successfully created with the intention that they belong to the same kind as current and previous X's (p. 10)'. That is, an object is a chair if it was successfully created with the intention that it would be a chair. Bloom (1996) further suggests that the way we infer the intention underlying a certain artifact kind is through its appearance and potential use. For instance, when we see a chair-shaped object, we infer that it was created to be a chair even when nobody tells us the intention of the designer who produced this object. Consequently, certain surface features become more important than others to the extent that they allow us to infer the designer's intentions. For instance, the overall shape of a chair would be essential but the color or the texture would be irrelevant. This account is quite compatible with the causal status hypothesis. As demonstrated in Experiment 2, physical features vary in their importance depending on their causal status; or, as in Bloom's theory, physical features related to underlying intentions would be considered more important, or essential.

Bloom (1996, 1998) also resorts to the intentional account in explaining why in Malt and Johnson (1992) some anomalous physical features did not affect membership likelihood (e.g. a small cube with an extendible rod used to register distances on a digital display was accepted as a ruler), whereas others did (e.g. a rubber sphere hitched to a team of dolphins was rejected as a boat): we can envision someone creating the cube with the intent to make a (more advanced, more effective) ruler, while it is implausible that someone who intended to build an object belonging to the class of boats would make the sphere (Bloom, 1998). (See Bloom, 1998 and Malt and Johnson, 1998 for more detail on this debate.) That is, it is the causal role of the atypical physical features that determines their acceptability.

On the other hand, the causal status hypothesis is critically different from Bloom's view for at least two reasons. First, the causal status hypothesis claims that any features, whether they are intentions or physical features, can be essential as long as

they are causally central. Second, the causal status hypothesis does not argue that a causally central feature is a defining feature as discussed earlier. Instead, cause features simply receive heavier weights than effect features. On the other hand, Bloom argues that intentions are necessary and sufficient for artifact categorization. In fact, Bloom discusses instances where intentions might not be necessary, as in a case in which lightning strikes a rock and shapes the rock into a chair. Clearly, no human intention of making a chair was involved here, but most people would call it a chair. Bloom argues, however, that when people see a complex design, they are drawn to the belief that it must have been created by some intentional force. The causal status hypothesis would account for the rock-chair-made-by-lightning example in a somewhat different way: although a cause (i.e. the intention) is missing (and therefore categorization as a chair might be less likely than when an effect feature is missing), the presence of all the other features of chairs would allow the object to pass the threshold of being called a chair. Again, this is because the causal status hypothesis does not assume that a cause feature is a defining feature.

4.3. Implications for categorization research

The current results deliver both good and bad news for existing categorization models and theories. The good news is that as far as feature weighting is concerned, the categorization models may not need to be concerned with domain differences. The distinction between natural kinds and artifacts has received a great deal of attention, partly because of its role in demonstrating the effect of domain theories. For instance, arbitrary categories such as 'white things' and natural kinds such as 'mammals' must be treated differently because our domain theories state that natural kinds have deep properties in common (Schwartz, 1979; Markman, 1989). These domain-specific effects could be troublesome in implementing categorization models for a number of reasons. It is not clear how different categorization models should be developed for each domain, and furthermore, it is not yet clear what different domains actually mean (Hirschfeld and Gelman, 1994). However, the current results indicate that weighting of features can operate in a domain-general manner. So far, the causal status effect has been demonstrated in the medical domains and social domains (Ahn et al., in preparation), as well as in natural kinds, artifacts, and nominal kinds. Therefore, it would be computationally feasible to add this domain-general characteristic onto any existing categorization model.

The bad news is that no existing model of categorization seems to be able to account for the causal status effect. Clearly, models assuming independence of features (e.g. Tversky, 1977; Anderson, 1991) cannot do so because the causal relations among features should be taken into account. Some models and theories are sensitive to correlated structures among features (e.g. Medin and Schaffer, 1978). However, causal relations are asymmetric, directional relations rather than symmetric correlations. Thus, these models, in simply being sensitive to correlations of features, cannot capture the differences within a correlation. Gentner (1989) has emphasized the importance of relational features as opposed to independent attri-

butes in inductive reasoning, but the current results show that it is not only the relational features, but also the relative status of given features within relations that matters.

4.4. Implications for developmental differences

The current interpretation of the experiments reported in this paper is consistent with developmental differences in categorization. Keil (1989) has shown that younger children were less likely to be impressed by discoveries of new internal properties in determining natural kind membership than were older children. Also, when external features were transformed (e.g. a raccoon being painted black with a single white stripe down the center of its back), younger children were much more likely to switch category membership (e.g. into a skunk in this example) than were older children. In both types of tasks, as children grew older, they became more affected by internal features when determining the category membership of natural kinds.

In the current framework, what may be developing is causal background knowledge about how internal features cause other surface features. That is, the indiscriminate nature of younger children's judgments might be because they have not yet acquired causal background knowledge about how types of features are related. It is still debatable whether this developmental difference derives from differences in domain-specific background knowledge about causal relationships, or from differences in a domain-general bias toward weighting cause more than effect. Given that even one-year-old children understand causality (Leslie and Keeble, 1987), and that even 10-month old infants pay more attention to causal agents than causal recipients (Cohen and Oakes, 1993), it is likely that younger children would also exhibit the bias of weighing cause more than effect. In this case, the non-discriminate responses in young children seem to be due simply to the lack of domain-specific causal knowledge. All these issues deserve future investigation.

5. Conclusion

The current study provides an explanation for domain differences, namely the category-feature interaction effect. The current study shows evidence against the belief that qualitatively different types of features must serve as defining or central features for natural kinds and artifacts. Instead, the apparent differences in the types of central features seem to derive from differences in the causal status of features.

Acknowledgements

I would like to thank Mary Lassaline for her help and discussion in the initial phase of this project. I also would like to thank two anonymous reviewers, Lewis Bott, Susan Gelman, Evan Heit, Douglas Medin, Steven Sloman, Neil Stewart, the

cognitive brown bag lunch group at University of Illinois Urbana-Champaign, and the cognitive brown bag lunch group at Yale University for their extremely helpful comments on this research. In addition, I would like to thank Helen Sullivan, Peter Jaros, and Yani Indrajana for collecting and coding the data, Marvin Chun for his help in programming the experiments, and Nancy Kim and Jennifer Amsterlaw for their thorough proofreading of the manuscript. Some of the stimulus materials used in Experiments 3 and 4 are adapted from the stimulus materials used in Rehder and Hastie (1997) and I would like to thank them for inspiring many of the features and objects used in these studies. This project was supported partly by a National Science Foundation Grant (NSF-SBR 9515085) and partly by a National Institute of Mental Health Grant (RO1 MH57737) given to the author.

Appendix A. Features used in Experiment 2

Category	Features
Boat	*Physical features* 1. Is wedge-shaped 2. Has a sail 3. Has an anchor 4. The sides are made of wood *Functional features*: carries people over water for purposes of work or recreation
Boot	*Physical features* 5. Extends above the ankle 6. Has a heel 7. Is made of leather *Functional features*: protects the feet from the elements: prevents excessive strain on the feet especially while hiking
Couch	*Physical features* 8. The frame is made of wood 9. Is covered with cloth 10. Filled with foam 11. Has cushions *Functional features*: seats 3–4 people comfortably
Desk	*Physical features* 12. The surface is flat 13. The surface is rectangular 14. Has drawers underneath 15. Has four legs 16. Has a matching chair *Functional features*: serves as a surface for studying, writing, or doing work in general; stores materials for studying, writing, or work
Paintbrush	*Physical features* 17. Has coarse bristles 18. The bristles are held by a metal band 19. The handle is made of wood 20. Has a handle *Functional features*: applies paint to surfaces, especially for smaller areas and corners

Rifle	*Physical features* 21. Has a barrel 22. The barrel is made of metal 23. Has an endpiece 24. The endpiece is made of wood 25. Has a trigger *Functional features*: shoots bullets one at a time from long range with great - accuracy without the need for reloading
Ruler	*Physical features* 26. Is 12 inches long 27. Is made of wood 28. Has a series of lines marked by numbers *Functional features*: measures distances of up to 1 foot; draws lines of a specific length up to 1 foot
Slacks	*Physical features* 29. Are made of polyester 30. Have a zipper 31. Have belt loops 32. Have cuffs *Functional features*: cover a person from the waist to the ankles often for formal occasions providing a nice appearance
Stove	*Physical features* 33. Has sides 34. The sides are made of sheet metal 35. Has burners 36. The number of burners is four 37. Has dials *Functional features*: cooks foods requiring careful attention by providing heat from underneath a pot or pan
Sweater	*Physical features* 38. Is made of wool 39. Has buttons down the front 40. Has sleeves 41. The sleeves end in small opening *Functional features*: provides extra warmth for the arms and upper body by being worn over a shirt
Taxi	*Physical features* 42. Has a meter for fares 43. Has two seats 44. Is painted yellow *Functional features*: provides private land travel for 1-4 people at a time when their own cars are unavailable and they are willing to pay the fare
Tractor	*Physical features* 45. Has an engine 46. Has large wheels 47. Has one seat 48. The seat is unenclosed 49. Has an attachment for farm machinery *Functional features*: allows one person on a farm to till ground or plow fields by pulling a variety of other machines

Appendix B. Appendix B. Natural kind and artifact versions of object descriptions and categorization questions used in Experiment 3

Note: (1) The order of feature presentation and question presentation was completely counterbalanced in Experiment 3. (2) C-Feature refers to compositional feature and F-feature refers to functional feature. C-Question refers to a question about an item that possesses only a compositional feature and F-Question refers to a question about an item that possesses only a functional feature. (3) The parts written in italics were not presented to the participants.

Appendix B.1.

1. *Natural kind version*: On the volcanic island of Kehoe, near Guam, there is a species of ant called Kehoe ants.

Compositional feature: 80% of Kehoe ants have blood high in iron sulphate.
Functional feature: 80% of Kehoe ants digest food fast.
C-Cause: blood high in iron sulphate tends to cause fast food digestion; the iron sulphate in Kehoe ants' blood stimulates the enzymes responsible for manufacturing the food-digesting secretions, and a Kehoe ant can digest food faster with more secretions.
F-Cause: Kehoe ants' fast food digestion tends to cause blood high in iron sulphate; fast food digestion increases Kehoe ants' metabolism which results in the blood high in iron sulphate.
C-Question: suppose an ant has blood high in iron sulphate, but does not digest food fast. How likely is it that this ant is a Kehoe ant?
F-Question: suppose an ant digests food fast, but does not have blood high in iron sulphate. How likely is it that this ant is a Kehoe ant?

2. *Artifact version*: On the volcanic island of Kehoe, near Guam, there is a kind of septic system called Kehoe septic.

Compositional feature: 80% of Kehoe septic systems have liquid high in iron sulphate.
Functional feature: 80% of Kehoe septic systems decompose sewage disposal fast.
C-Cause: liquid high in iron sulphate tends to cause fast decomposition of sewage disposal; The iron sulphate stimulates anaerobic bacteria responsible for decomposition, and a Kehoe septic system can decompose disposals faster with more bacteria.
F-Cause: fast decomposition of sewage disposal tends to cause liquid high in iron sulphate; The anaerobic bacteria responsible for fast decomposition produce iron sulphate as a by-product while decomposing disposals.

C-Question: suppose a septic system has liquid high in iron sulphate, but does not decompose sewage disposal fast. How likely is it that this septic system is a Kehoe septic system?
F-Question: suppose a septic system decomposes sewage disposal fast, but does not have liquid high in iron sulphate. How likely is it that this septic system is a Kehoe septic system?

Appendix B.2. Coryanthes

1. *Natural kind version*: in Central Africa, there are orchids called 'Coryanthes'.

Compositional feature: 80% of Coryanthes have a component called 'Eucalyptol' in their flowers.
Functional feature: 80% of Coryanthes attract orchid male bees.
C-Cause: because Coryantheses usually have Eucalyptol, Coryantheses tend to attract orchid male bees; Eucalyptol which has a strong smell of eucalyptus tend to be attractive to most male orchid bees.
F-Cause: because Coryantheses tend to attract orchid male bees, they usually have Eucalyptol in their flowers; While orchid male bees are collecting nectar from Coryantheses, they tend to leave Eucalyptol obtained from other flowers in Coryantheses.
C-Question: suppose a flower in Central Africa has Eucalyptol in the flower, but does not attract orchid male bees. How likely is it that this flower is a Coryanthes?
F-Question: suppose a flower in Central Africa attracts orchid male bees, but does not have Eucalyptol in the flower. How likely is it that this flower is a Coryanthes?

2. *Artifact version*: In Central Africa, there are pillars called 'Coryanthes'.

Compositional feature: 80% of Coryanthes have a component called 'Eucalyptol' in their top.
Functional feature: 80% of Coryanthes attract orchid male bees.
C-Cause: because Coryantheses are usually built with Eucalyptol, they tend to attract orchid male bees; Eucalyptol which has a strong smell of eucalyptus tend to be attractive to male orchid bees.
F-Cause: because Coryantheses attract orchid male bees, they tend to have Eucalyptol in their top; While orchid male bees are storing collected nectar in the top of Coryantheses, they tend to leave Eucalyptol obtained from other flowers in Coryantheses.
C-Question: suppose a pillar in Central Africa attracts orchid male bees, but does not have Eucalyptol in the top. How likely is it that this barn is a Coryanthes?
F-Question: suppose a pillar in Central Africa has Eucalyptol in the top, but does not attract orchid male bees. How likely is it that this barn is a Coryanthes?

Appendix B.3. Yabuka

1. *Natural kind version*: In Northern Alaska, there are rocks called 'Yabuka'.

Compositional feature: 80% of Yabukas contain carbon.
Functional feature: 80% of Yabukas are used as fuel for heaters.
C-Cause: because Yabukas tend to contain carbon, they tend to be used as fuel for heaters; Yabukas were not initially used as fuel. But the Alaskans discovered that carbon in Yabukas burns at a hot temperature.
F-Cause: because Yabukas are used as fuel for heaters, they tend to contain carbon; Yabukas initially don't have carbon. When Yabukas burn, carbon tends to be created in Yabukas.
C-Question: suppose a rock contains carbon, but has not been used as fuel for heaters. How likely is it that this rock is a Yabuka?
F-Question: suppose a rock is used as fuel for heaters, but does not contain carbon. How likely is it that this rock is a Yabuka?

2 *Artifact version*: Northern Alaskans have invented synthesized materials called 'Yabuka'.

Compositional feature: 80% of Yabukas contain carbon.
Functional feature: 80% of Yabukas are used as fuel for heaters.
C-Cause: because Yabukas contain carbon, they tend to be used as fuel for heaters; Yabukas were initially used as building materials. But the Alaskans discovered that carbon in Yabukas burns at a hot temperature.
F-Cause: because Yabukas are used as fuel for heaters, they tend to contain carbon; Yabukas initially didn't have carbon. When Yabukas burn, carbon tends to be created in Yabukas.
C-Question: suppose a synthesized material contains carbon, but has not been used as fuel for heaters. How likely is it that this synthesized material is a Yabuka?
F-Question: suppose a synthesized material is used as fuel for heaters, but does not contain carbon. How likely is it that this synthesized material is a Yabuka?

Appendix B.4. Oenothera

1. *Natural kind version*: There are plants called 'oenotheras' on a Pacific island.

Compositional feature: 80% of Oenotheras have lots of protein in their stems.
Functional feature: 80% of Oenotheras are used to kill insects.
C-Cause: because Oenotheras usually have lots of protein in their stems, they tend to be used to kill insects; Oenotheras were not initially used by humans. Since oenotheras eat anything that touch them and the smell of the protein in Oenotheras' stems attract insects, oenotheras tend to be used to kill insects.
F-Cause: because Oenotheras tend to be used to kill insects, they usually have lots of protein in their stems; Oenotheras initially didn't have protein in their stems. When Oenotheras are used to kill insects, Oenotheras eat many insects and the excessive protein tends to be built up in Oenotheras' stems as a

result.

C-Question: suppose a plant on the island has lots of protein in its stem but is never used to kill insects. How likely is it that this plant is an oenothera?

F-Question: suppose a plant on the island is used to kill insects but does not have protein in its stem. How likely is it that this plant is an oenothera?

2. *Artifact version*: People in a pacific island invented a machine called 'oenothera'.

Compositional feature: 80% of Oenotheras have lots of protein built up in their container.

Functional feature: 80% of Oenotheras are used to kill insects.

C-Cause: because Oenotheras tend to have lots of protein built-up in their container, they tend to be used to kill insects; Oenotheras are fertilizer dispensers. The smell of fertilizer protein built up in the container draws insects, which then get killed while fertilizer is dispensed.

F-Cause: because oenotheras tend to be used to kill insects, they tend to have lots of protein built up in its container; Oenotheras' container initially didn't have protein built up. As Oenotheras kill insects, they tend to use the protein from the insects to strengthen the walls of the container.

C-Question: suppose a machine has protein built up in its container but is not used to kill insects. How likely is it that this machine is an oenothera?

F-Question: suppose a machine is used to kill insects but does not have protein built up in its container. How likely is it that this machine is an oenothera?

Appendix C. Stimulus materials used in Experiment 4

Note: (1) The order of feature presentation and question presentation was completely counterbalanced in Experiment 4. (2) C-Feature refers to compositional feature and F-feature refers to functional feature. C-Question refers to a question about an item that possesses only a compositional feature and F-Question refers to a question about an item that possesses only a functional feature. (3) The parts written in italics were not presented to the participants.

Appendix C.1. Natural kind 1

Lake Victoria Shrimp are found in Lake Victoria, Africa.80% of Lake Victoria shrimps have a high quantity of ACh neurotransmitter.80% of Lake Victoria shrimps show a long-lasting flight response.

F-Cause: a long-lasting flight response tends to cause a high quantity of ACh neurotransmitter; the greater flight response consumes more energy, increasing the amount of ACh neurotransmitter.

C-Cause: a high quantity of ACh neurotransmitter tends to cause a long-lasting flight response; The duration of the electrical signal to the muscles is longer because of the excess amount of neurotransmitter.

C-Question: suppose a shrimp has a high quantity of ACh neurotransmitter but does not show a long-lasting flight response. How likely is it that this shrimp is a

Lake Victoria Shrimp?
F-Question: suppose a shrimp shows a long-lasting flight response, but does not have a high quantity of ACh neurotransmitter. How likely is it that this shrimp is a Lake Victoria Shrimp?

Appendix C.2. Natural kind 2

There are fruits called 'nidems' that grow in a Northern African country.80% of Nidems have firm skin.80% of Nidems are used as a wheel for carts.

C-Cause: Nidems' firm skins allow them to be frequently used as wheels for carts; firm skins of nidem fruits are suitable as wheels because they last long.
F-Cause: using nidems as a wheel for carts tends to make their skin firm; the fiber in the skin of nidem fruits are initially soft but it becomes firm as the fruits are rolled on the road as wheels.
C-Question: suppose you are in this country and see a fruit with very firm skin, but it has never been used as a wheel for carts. How likely is it that this fruit is a nidem?
F-Question: suppose you are in this country and see a fruit that is used as a wheel for carts, but it does not have very firm skin. How likely is it that this fruit is a nidem?

Appendix C.3. Natural kind 3

In eastern Asia, there is a type of flower called 'phyrum'.80% of phyrum flowers have velvety leaves.80% of phyrum flowers are used to repel mosquitoes.

C-Cause: because phyrum flowers have velvety leaves, they tend to be used as a mosquito repellent; Velvety leaves naturally repels mosquitoes by reflecting dim light in a peculiar way.
F-Cause: because phyrum flowers are used to repel mosquitoes, their leaves tend to become velvety; Often animals and humans rub themselves against the phyrum leaves in order to repel mosquitoes. As phyrum leaves are rubbed, their leaves, which were initially rough, become velvety.
C-Question: suppose you see a flower in eastern Asia and this flower has velvety leaves but is never used as a mosquito repellent. How likely is it that this is a phyrum flower?
F-Question: suppose you see a flower in eastern Asia and this flower is used as a mosquito repellent but does not have velvety leaves. How likely is it that this is a phyrum flower?

Appendix C.4. Artifact 1

Neptune personal computer are made by the military.80% of Neptune computers have a hot computer temperature.80% of Neptune computers display a bright screen

image.

C-Cause: hot temperature tends to make Neptune computers display a bright image; heat increases the efficiency of the cathode ray tube, leading to a more energized electron beam and a brighter screen.
F-Cause: a bright screen image in Neptune computers tends to cause hot computer temperature; displaying a bright screen image requires fast and intensive use of components in the cathode ray tube, making the components heated up.
F-Question: Suppose a computer displays a bright screen image but does not have hot temperature. How likely is it that this computer is a Neptune computer?
C-Question: Suppose a computer has hot temperature but does not display a bright screen image. How likely is it that this computer is a Neptune computer?

Appendix C.5. Artifact 2

A chemist in northern Italy created an artificial flavor called tansline.80% of Tansline has a tangy taste.80% of Tansline is used for fruit tarts.

C-Cause: Tansline's tangy taste allows it to be frequently used for fruit tarts; Northern Italians are fond of fruit tarts that taste tangy.
F-Cause: Using Tansline for fruit tarts tends to cause it to taste tangy; When Tansline is used for making fruit tarts, a special chemical reaction occurring during the cooking process causes tansline to have a tangy taste.
F-Question: Suppose you are in northern Italy and taste an artificial flavor that was used for making fruit tarts but it does not taste tangy. How likely is it that this artificial flavor is tansline?
C-Question: Suppose you are in northern Italy and taste an artificial flavor that is tangy but was never used for making fruit tarts. How likely is it that this artificial flavor is tansline?

Appendix C.6. Artifact 3

In Australia, there is a machine called Kankakin.80% of Kankakins have a rubber platform.80% of Kankakins are used for relaxing pregnant mares during delivery.

C-Cause: because Kankakins have a rubber platform, they tend to be used for relaxing pregnant mares; a Kankakin has a vibrating rubber platform and is used on fishing boats to sort shell fish. Horse breeders found that Kankakin machines can be used during a mare's labor because the gentle vibration of the Kankakin machine's rubber platform soothes the mare during labor.
F-Cause: because Kankakins are used relaxing pregnant mares, they tend to have a rubber platform; Kankakins vibrate during a mare's labor to soothe the mare. In order to make vibration more soothing, they changed the machine's wooden platform into a rubber platform.
F-Question: suppose you see a machine in Australia and this machine is used for

relaxing pregnant mares during delivery but it does not have a rubber platform. How likely is it that this machine is a Kankakin?

C-Question: suppose you see a machine in Australia and this machine has a rubber platform but has never been used for relaxing pregnant mares during delivery. How likely is it that this machine is a Kankakin?

References

Ahn, W., Brewer, W.F., Mooney, R.J., 1992. Schema acquisition from a single example. Journal of Experimental Psychology: Learning, Memory and Cognition 18 (2), 391–412.

Ahn, W., Lassaline, M.E., 1995. Causal structure in categorization. Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society, Pittsburgh, PA, pp. 521–526.

Anderson, J.R., 1991. The adaptive nature of human categorization. Psychological Review 98 (3), 409–429.

Atran, S., 1987. Origin of the species and genus concepts: an anthropological perspective. Journal of the History of Biology 20, 195–279.

Barr, R.A., Caplan, L.J., 1987. Category representations and their implications for category structure. Memory and Cognition 15, 397–418.

Barton, M.E., Komatsu, L.K., 1989. Defining features of natural kinds and artifacts. Journal of Psycholinguistic Research 18, 433–447.

Billman, D., 1989. Systems of correlations in rule and category learning: use of structured input in learning syntactic categories. Language and Cognitive Processes 4, 127–155.

Billman, D., Knutson, J., 1996. Unsupervised concept learning and value systematicity: a complex whole aids learning the parts. Journal of Experimental Psychology: Learning, Memory and Cognition 22, 458–475.

Bloom, P., 1996. Intention, history, and artifact concepts. Cognition 60, 1–29.

Bloom, P., 1998. Theories of artifact categorization. Cognition (in press).

Braisby, N., Franks, B., Hampton, J., 1996. Essentialism, word use, and concepts. Cognition 59, 247–274.

Bruner, J., Goodnow, J., Austin, G., 1978. A study of thinking. In: Langfeld, H. (Ed.), A Wiley Publication in Psychology. Wiley, New York, pp. 1–22.

Carey, S., 1985. Conceptual Change in Childhood. Plenum, Cambridge, MA.

Cohen, L.B., Oakes, L.M., 1993. How infants perceive a simple causal event. Developmental Psychology 29, 421–433.

Cohen, J.D., MacWhinney, B., Flatt, M., Provost, J., 1993. Psyscope: a new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments and Computers 25 (2), 257–271.

Gelman, S.A., 1988. The development of induction within natural kind and artifact categories. Cognitive Psychology 20, 65–95.

Gelman, S.A., Coley, J.C., Gottfried, G.M., 1994. Essentialist beliefs in children. In: Hirschfeld, L.A., Gelman, S.A. (Eds.), Mapping the Mind: Domain Specificity in Cognition and Culture. Cambridge University Press, Cambridge, MA, pp. 341–365.

Gelman, S.A. and Kalish, C.W., 1993. Categories and causality. In: Pasnak, R., Howe, M.L. (Eds.), Emerging Themes in Cognitive Development, Vol. 2. Springer, New York.

Gelman, S.A., Wellman, H.M., 1991. Insides and essences: early understandings of the nonobvious. Cognition 38, 213–244.

Gentner, D., 1989. The mechanisms of analogical learning. In: Vosniadou, S., Ortony, A. (Eds.), Similarity and Analogical Reasoning. Cambridge University Press, Cambridge.

Hampton, J.A., 1995. Testing the prototype theory of concepts. Journal of Memory and Language 34, 686–708.

Hirschfeld, L.A., Gelman, S.A., 1994. Mapping the Mind: Domain Specificity in Cognition and Culture. Cambridge University Press, Cambridge, MA.

Hunt, S.M.J., 1994. MacProbe: a Macintosh-based experimenter's workstation for the cognitive sciences. Behavior Research Methods, Instruments and Computers 26, 345–351.

Kalish, C., 1998. Natural and artifactual kinds: are children realists or relativists about categories? Developmental Psychology 34.

Keil, F.C., 1989. Concepts, Kinds, and Cognitive Development. MIT Press, Cambridge, MA.

Kripke, S., 1971. Naming and necessity. In: Davidson, D., Harman (Eds.), Semantics of Natural Language. Reidel, Dordrecht.

Leslie, A.M., Keeble, S., 1987. Do six-month-old infants perceive causality? Cognition 25, 265–288.

Locke, J., 1894/1975. An Essay Concerning Human Understanding. Oxford University Press, Oxford.

Malt, B.C., 1994. Water is not H_2O. Cognitive Psychology 27, 41–70.

Malt, B.C., Johnson, E.C., 1992. Do artifact concepts have cores? Journal of Memory and Language 31, 195–217.

Malt, B.C. and Johnson, E.C., 1998. Artifact Category Membership and the Intentional-Historical Theory. Cognition.

Malt, B.C., Smith, E.E., 1984. Correlated properties in natural categories. Journal of Verbal Learning and Verbal Behavior 23, 250–269.

Markman, E.M., 1989. Categorization and Naming in Children: Problems of Induction. MIT Press, Cambridge, MA.

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

Medin, D.L., Ortony, A., 1989. Psychological essentialism. In: Vosniadou, S., Ortony, A. (Eds.), Similarity and Analogical Reasoning. Cambridge University Press, Cambridge, MA, pp. 179–195.

Medin, D.L., Schaffer, M.M., 1978. Context theory of classification learning. Psychological Review 85, 207–238.

Murphy, G.L., 1993. A rational theory of concepts. In: Nakamura, G.V., Taraban, R., Medin, D.L. (Eds.), The Psychology of Learning and Motivation. Academic Press, San Diego, CA, pp. 327–359.

Murphy, G.L., Medin, D.L., 1985. The role of theories in conceptual coherence. Psychological Review 92, 289–316.

Nosofsky, R.M., Palmeri, T.J., McKinley, S.C., 1994. Rule-plus-exception model of classification learning. Psychological Review 101 (1), 53–79.

Pazzani, M.J., 1991. Influence of prior knowledge on concept acquisition: experimental and computational results. Journal of Experimental Psychology: Learning, Memory and Cognition 17 (3), 416–432.

Putnam, H., 1975. The meaning of ''meaning''. In: Mind, Language and Reality, Vol. 2: Philosophical papers. Cambridge University Press, Cambridge, MA.

Putnam, H., 1977. Is semantics possible? In: Schwartz, S.P. (Ed.), Naming, Necessity, and Natural Kinds. Cornell University Press, Ithaca, NY.

Rehder, B., Hastie, R., 1997. The roles of causes and effects in categorization. Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 650–655.

Rips, L.J., 1989. Similarity, typicality, and categorization. In: Stella Vosniadou, A.O. (Ed.), Similarity and Analogical Reasoning. Cambridge University Press, New York, pp. 21–59.

Rosch, 1978. Principles of categorization. In: Rosch, E., Lloyd, B.B. (Eds.), Cognition and Categorization. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 27–47.

Schwartz, S.P., 1979. Natural kind terms. Cognition 7, 301–315.

Sloman, S.A., Ahn, W., 1998. Feature centrality: naming versus imaging. Memory & Cognition, in press.

Sloman, S.A., Love, B.C., Ahn, W., 1998. Feature centrality and conceptual coherence. Cognitive Science 22, 189–228.

Tversky, A., 1977. Features of similarity. Psychological Review 84, 327–352.

Wattenmaker, W.D., 1995. Knowledge structures and linear separability: integrating information in object and social categorization. Cognitive Psychology 28, 274–328.