UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Inferring Unobserved Category Features With Causal Knowledge

Permalink

Journal

ISSN

Authors

Rehder, Bob
Burnett, Russell C

Publication Date

2002

Peer reviewed

Inferring Unobserved Category Features With Causal Knowledge

Bob Rehder (**bob.rehder@nyu.edu**)
Department of Psychology, New York University, 6 Washington Place
New York, NY 10003 USA

Russell C. Burnett (**r-burnett@northwestern.edu**)
Department of Psychology, Northwestern University, 2029 Sheridan Road
Evanston, IL 60208 USA

Abstract

One central function of categories is to allow people to in-
fer the presence of features that cannot be directly ob-
served. Although the effect of observing past category
members on such inferences has been considered, the effect
of theoretical or causal knowledge about the category has
not. We compared the effects of causal laws on feature pre-
diction with the effects of the inter-feature correlations that
are produced by those laws, and with the effect of exemplar
typicality or similarity. Feature predictions were strongly
influenced by causal knowledge. However, they were also
influenced by similarity, in violation of normative behav-
ior as defined by a Bayesian network view of causal reason-
ing. Finally, feature predictions were not influenced by the
presence of correlations among features in observed cate-
gory members, indicating that causal relations versus cor-
relations lead to different inferences regarding the presence
of unobserved features.

When an object has been classified as an instance of a con-
cept, knowledge associated with that concept can be brought
to bear in reasoning about the features that the object is
likely to possess. But what is the nature of that knowledge,
and how is it used to make inferences or predictions about
unobserved features? Recent research has demonstrated that
tasks such as category learning, categorization, and category-
based induction are often influenced by the theoretical
knowledge that one possesses. This knowledge often takes
the form of causal relations between features of a category,
and theories have been proposed to account for the effects of
such knowledge (Rehder, 1999, 2001; Waldmann, Holyoak,
& Fratianne, 1995). In this article we assess the effect of
causal relations on feature inferences, and in the first of the
following sections we present a formal model of causal
knowledge and its predictions regarding feature inferences.

Of course, another form of knowledge that may guide fea-
ture inference is empirical information derived from the first-
hand observation of category members. Prior research sug-
gests two likely effects of such empirical knowledge on fea-
ture prediction. First, feature predictions will often be influ-
enced by the overall similarity to the category of the exem-
plar with the unobserved feature. In the second section we
discuss this predicted effect of similarity and show how it
can run directly counter to the predictions of our formal
model of causal knowledge. Second, the presence of correla-
tions among category features may also allow one to infer

the presence of a feature given knowledge about the presence
of one or more other features. We discuss the effects of ob-
served inter-feature correlations in the third section, and
compare them to the effects produced by direct knowledge of
causal relations—relations that were responsible for generat-
ing the feature correlations in the first place.

Feature Inference via Causal Reasoning

It is clear that causal knowledge has predictive value. For
example, given knowledge of the causes of fire, one can
predict, with some certainty, that a flame will appear when a
match is struck, oxygen is present, and so on. Likewise,
given the causal relations that hold among features of an
object, the presence of an unobserved feature can be inferred
by reasoning about the causes of that feature and whether
those causes are present in the object at hand.

In this article we provide direct evidence of causal reason-
ing in feature inference, and we test a well specified theory
about how this sort of reasoning might be done. This theory
involves Bayesian networks—graphs in which variables are
represented as nodes, and causal relations between the vari-
ables as directed links between the nodes. Figure 1 shows a
simple Bayesian network in which three effect variables are
dependent on a single cause variable.

Rules by which inferences can be drawn from Bayesian
networks have been well developed in artificial intelligence.
One important rule is the *causal Markov condition*, which
states that a variable X is independent of all variables that
are not themselves descendents of X given knowledge about
the state of X's (immediate) parents (Pearl, 1988). In Figure
1, for example, the state of F_2 is independent of F_3 and F_4
given knowledge about F_1.

It has been proposed that Bayesian networks are good psy-
chological models of causal knowledge—and, in particular,
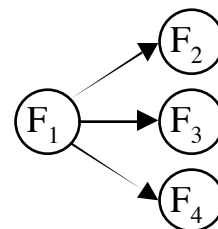of the causal knowledge associated with object concepts

Figure 1. A common-cause causal schema.

(Rehder, 1999, 2001; Waldmann et al., 1995). On this *causal-model theory* of concepts, the model shown in Figure 1 can be used to represent a concept with four features, where feature F_1 causes features F_2, F_3, and F_4. Representing causal knowledge of category features in this way has been shown to account well for classification (Rehder, 1999, 2001), but it is an open question whether the rules of inference associated with Bayesian networks—and, in particular, the causal Markov condition—accurately describe people's inferences about unobserved features.

In this article we tackle this question by explicitly manipulating the causal knowledge about a category that a participant has and measuring the effects of that knowledge on subsequent feature predictions. Participants in a *Causal Schema* condition were told that the features of a novel category were related as in the common-cause model of Figure 1, and were then asked to make inferences about exemplars in which one of the four features was unobservable. We determined whether participants' responses were consistent with the causal Markov condition.

Feature Inference via Category Similarity

Another factor likely to influence the prediction of an unobserved feature is the similarity of the exemplar to previously observed category members (or to the category's prototype). In this case, inference is based on simple feature overlap. If an exemplar is similar to (i.e., shares many features with) many category members, and if many category members possess the unobserved feature, then the exemplar probably has the feature too. For example, if a new bird has many features typical of birds (small, sings, eats worms, etc.) it probably also flies, because it is similar to many birds and most birds fly. In contrast, a new bird with many atypical features (e.g., an ostrich) is similar to fewer birds, and so the inference to flight is less certain.

Previous research has shown that similarity plays a key role in a variety of feature inference tasks. For example, Sloman's (1993) feature-based model of the inductive projection of features across categories assumes that a feature is projected from, say, robins to falcons by computing the extent to which they have other features in common (cf. Rips, 1975). Sloman also found a phenomenon called "inclusion similarity" in which participants projected a property from an inclusive category to a subordinate (e.g., from bird to robin) more strongly when the subordinate was more typical (e.g., robin) than when it was less so (e.g., penguin). Direct evidence of the role of similarity in feature prediction (rather than projecting new features across categories) was provided by Yamauchi and Markman (2000), who taught participants artificial categories and found that exemplars that were closer to the category prototype (i.e., that possessed more features in common with training exemplars) supported stronger inferences of unobserved features.

The influence of similarity on feature inference presents a particularly stringent test of the causal Markov condition, because honoring the causal Markov condition can require one to ignore similarity. For a category with a common-cause causal schema (Figure 1), the causal Markov condition states that information about the presence or absence of F_2 and F_3 is irrelevant to inferring F_4 given knowledge of F_1. In contrast, an influence of similarity predicts that inferences to F_4 will be stronger when F_2 and F_3 are present, because the presence of F_2 and F_3 means that the exemplar is more similar to the category prototype.

In the following experiment participants in the *Control* condition were told that each feature had a 75% base rate (as were Causal Schema participants), but were not instructed on any causal relationships. Results from the Control condition will indicate an effect of similarity if feature inferences increase as a function of the exemplar's similarity to this central tendency (i.e., as a function of the number of features). Results from the Causal Schema condition will indicate whether participants are able to override this effect of similarity, as required by the causal Markov condition.

Feature Inference via Feature Correlations

The final influence on feature prediction performance we consider is the presence of within-category feature correlations. For example, many people know that birds that are small tend to sing whereas large birds do not, and on the basis of this correlation might infer the presence of a small bird upon hearing song, or the absence of singing from a large bird—and do so despite having no knowledge of the causal mechanisms that link size and singing.

Prior research confirms the intuition that the observation of within-category feature correlations can influence feature predictions, at least when participants observe category exemplars during standard classification-with-feedback training. Some studies (Thomas, 1998; Yamauchi & Markman 1998) have attributed this result to participants' similarity-matching to the training exemplars with a multiplicative similarity rule that preserves sensitivity to feature correlations. Others (Anderson & Fincham, 1996) attribute it to participants' inducing a direct representation of those inter-feature correlations (also see Wattenmaker, 1993).

A final goal of the current article was to compare the effect of causal laws on feature inference with the effect of observing the inter-feature correlations produced by those laws. In the following experiment, participants in the *Exemplars* condition were told that each feature manifested a 75% base rate, as were participants in the Causal Schema and Control conditions. But then, rather than being instructed on causal relationships, they instead observed a sample of exemplars that manifested the inter-feature correlational structure that is implied by a common-cause causal schema (i.e., exemplars with strong correlations between feature F_1 and features F_2, F_3, and F_4). Because it reflects causal laws, feature prediction performance based on this correlational structure should ideally be qualitatively similar to performance based on knowledge of the laws alone. In particular, inferences regarding the presence of an unobserved effect feature should be stronger when F_1 is present as compared to one of the other features.

Method

Materials

Six novel categories were used: two biological kinds (Kehoe Ants, Lake Victoria Shrimp), two nonliving natural kinds

(Myastars, Meteoric Sodium Carbonate), and two artifacts (Romanian Rogos [cars], Neptune Personal Computers). Each category had four binary features. For example, for the Lake Victoria Shrimp category the four binary features were "a high quantity of ACh neurotransmitter," "long-lasting flight response," "accelerated sleep cycle," and "high body weight." Each feature was described as occurring in 75% of category members. Participants in the Causal Schema condition were also taught about three causal relationships between F_1 and F_2, F_3, and F_4. Each description of a causal relationship specified the cause feature, the effect feature, and a brief description of causal mechanism linking them. For example, the $F_1 \rightarrow F_2$ causal relationship for Lake Victoria Shrimp was "A high quantity of ACh neurotransmitter causes a long-lasting flight response. The duration of the electrical signal to the muscles is longer because of the excess amount of neurotransmitter."

Participants

Fifty-four undergraduates or other members of the Northwestern University community received course credit or pay for participating in this experiment.

Design

Participants were randomly assigned in equal numbers to one of the six categories, and to either the Causal Schema, the Exemplars, or the Control condition.

Procedure

All phases of the experiment were conducted by computer. Participants first studied several screens of information about the assigned category at their own pace. All participants read a cover story and a description of the features and their 75% base rates. Participants in the Causal Schema condition also received a description of three causal relationships, and a diagram depicting those relationships similar to Figure 1. When ready, all participants took a multiple-choice test of this knowledge. Participants could request help, which led the computer to re-present the information about the category. Participants were required to retake the test until they made 0 errors and 0 requests for help.

Participants in the Exemplars condition then observed 48 examples of the category. Although the studies reviewed above found feature prediction performance to be sensitive to feature correlations when training exemplars were observed in a classification-with-feedback task, Wattenmaker (1991) found that participants were more sensitive to feature correlations on a transfer categorization test when they were asked simply to "look over, examine, and learn about" exemplars. Thus, category exemplars were presented sequentially at a pace determined by the participants. They observed 26, 3, 3, 3, 1, 1, 1, 2, 2, 2, and 4 instances of exemplars 1111, 1110, 1101, 1011, 0110, 0101, 0011, 0100, 0010, 0001 and 0000, respectively, where "1" denotes the presence of a feature, "0" represents its absence, and features are given in dimension order (F_1, F_2, F_3, F_4). These exemplars manifest the 75% feature base rates that participants were instructed on, and also the correlational structure that is implied by a common-cause causal schema. Specifically, the strength of

the correlations between F_1 and F_2, F_3, and F_4 was $r = .62$, and the correlations among F_2, F_3, and F_4 conditional on F_1 were approximately 0. The features of each exemplar were listed in order (1–4) on the computer screen. For example, participants assigned to the Lake Victoria Shrimp category were presented with three category members that possessed "high amounts of the ACh neurotransmitter," "a normal flight response," "accelerated sleep cycle," and "high body weight." The order of the 48 exemplars was randomized for each participant.

Participants in all conditions then performed two tasks (counterbalanced for order): a feature prediction task and a categorization task. During the feature prediction task, participants were presented with 32 exemplars, each with an unobserved value on one of the four dimensions, and were asked to rate the likelihood that the feature was present on a 100-point scale. For each unobserved dimension the other three dimensions took on the eight possible combinations of values, yielding a total of 32 feature prediction problems. The features of each exemplar were listed in dimension order (1–4), with the unknown dimension designated with "???." For example, participants assigned to the Lake Victoria Shrimp category were presented with the feature list "normal amounts of the ACh neurotransmitter," "a fast flight response," "???," and "high body weight" and asked to rate on a 100-point scale whether this exemplar had an "accelerated sleep cycle." The order of the 32 feature prediction problems was randomized for each participant.

During the categorization task, participants rated the category membership of exemplars on a 100-point scale. There were 32 exemplars, consisting of all possible 16 examples that could be formed from four binary features, each presented twice. The order of the 32 test exemplars was randomized for each participant.

Results

Feature Prediction Results

Because results for those feature prediction problems in which the unobserved feature dimension was the first dimension were not directly relevant to the theoretical issues raised in this article, we report results only for those problems in which the unobserved feature was on the second, third, or fourth dimension. Figure 2 presents feature prediction ratings as a function of the total number of features in the exemplar, whether the common-cause feature (F_1) is present or absent in that exemplar, and experimental condition (Causal Schema, Control, or Exemplar). Note that the number of features in F_1-Present problems ranges from 1 to 3 whereas the number in F_1-Absent problems ranges from 0 to 2 because of the presence of F_1 itself in the F_1-Present problems.

In the Causal Schema condition (Figure 2a) feature prediction ratings were strongly influenced by the presence or absence of the common cause F_1 as compared to one of the other features. For example, problems with one feature received a much higher rating when that feature was F_1 (e.g., the feature prediction problem 1x00) than when it was one of the other features (e.g., 0x10), 70.6 versus 24.5. Similarly, problems with two features received a higher rating when one of those features was F_1 (e.g., 1x10) than when
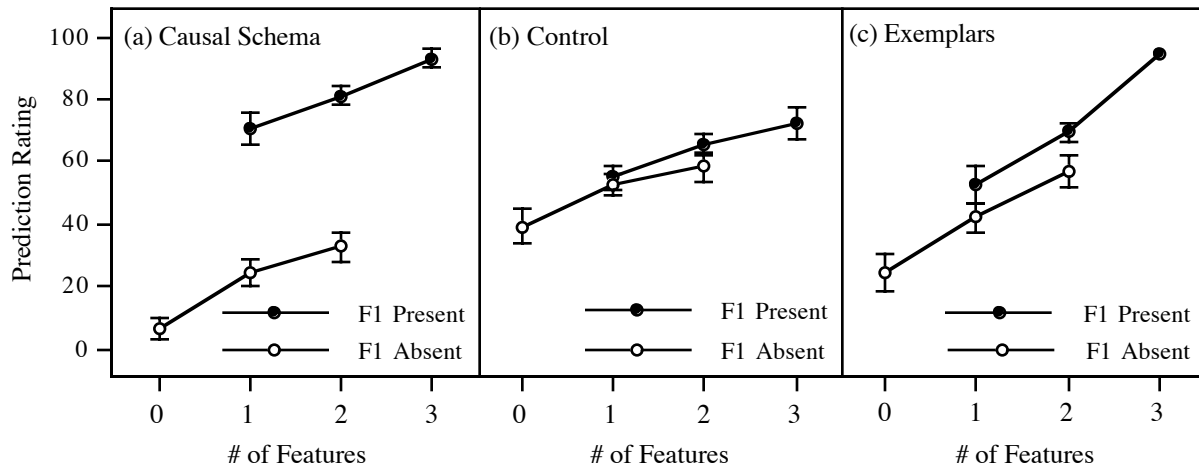
Figure 2. Feature prediction results in the (a) Causal Schema, (b) Control, and (c) Exemplars conditions

neither was F_1 (e.g., 0x11), 80.9 versus 32.8. That is, Causal Schema participants were much more likely to reason from the presence (or absence) of the common-cause feature F_1 to infer the presence (or absence) of an effect than when they were reasoning from one of the effect features.

According to the causal Markov condition, inferring the presence of an effect in a common-cause schema should not only depend on the common cause feature F_1, it should also *not* depend on any of the effect features. In fact, Figure 2a indicates that feature prediction ratings increased as the number of effect features increased. In the F_1-absent condition, feature prediction ratings were 6.7, 24.5, and 32.9 for exemplars possessing 0, 1, and 2 effect features, respectively. This occurred despite the fact that, according to the causal Markov condition, the absence of common cause F_1 makes the presence or absence of other effects irrelevant for predicting an effect feature. Likewise, in the F_1-present condition, ratings were 70.5, 80.9, and 92.8 for exemplars possessing 1, 2, and 3 features, respectively. That is, although Causal Schema participants' feature prediction ratings were strongly influenced by the causal knowledge that was provided, they also exhibited a substantial similarity effect in which more features led to stronger inferences, in violation of the causal Markov condition.

In comparison with the Causal Schema condition, results from the Control condition (Figure 2b) indicate that the effect of the presence or absence of F_1 on feature prediction ratings was not greater than the effects of the other features. This result was expected, because in the Control condition there was nothing about F_1 to make it especially predictive of an unobserved feature. However, as in the Causal Schema condition, feature prediction ratings exhibited an effect of similarity; ratings increased as a function of the number of features present in the exemplar. Ratings were 37.4, 53.6, 63.1, and 73.7 for exemplars that possessed 0, 1, 2, or 3 features, respectively. Note that this effect of similarity obtained despite the fact that the Control participants observed no members of the category, but rather just read a verbal statement of the 75% feature base rates.

These conclusions were supported by statistical analysis. Each participant's ratings were predicted from a regression

equation in which the two predictors were the number of features present and a contrast code representing the presence or absence of F_1. As expected, in the Causal Schema condition the regression weight associated with the presence or absence of F_1 was both significantly greater than zero, $t(35) = 6.95$, $p < .0001$, and significantly different than the corresponding weight in the Control condition, $t(34) = 5.65$, $p < .0001$. Moreover, in both the Causal Schema condition and the Control condition the regression weight associated with the number of features was significantly different from zero, $t(35) = 4.89$, $p < .0001$, and $t(35) = 2.79$, $p < .01$, respectively. This sensitivity to number of features did not differ between the Causal Schema and Control conditions, $t < 1$.

Finally, Figure 2c presents the results from the Exemplars condition. The figure indicates that, in contrast to the Causal Schema condition, the presence of F_1 resulted in only a small increase in feature prediction ratings as compared to one of the other features. For example, whereas in the Causal Schema condition problems received a feature prediction rating that was about 50 points higher when F_1 was present in the exemplar (Figure 2a), in the Exemplars condition that difference was only 10.9 points (52.9 vs. 42.0) for one-feature exemplars and 12.4 points (69.3 vs. 56.9) for two-feature exemplars. This result obtained despite the presence of strong correlations between F_1 and the other features in the training exemplars that might have been expected to lead subjects to treat F_1 as especially predictive. In fact, the regression weight associated with the presence or absence of F_1 in the Exemplars condition was not significantly different from the Control condition, $t(34) = 1.17$, $p > .20$.

As in the Causal Schema and Control conditions, feature prediction ratings in the Exemplars condition were sensitive to the total number of features possessed by the exemplar. Ratings were 24.2, 44.2, 65.1, and 94.0 for exemplars that possessed 0, 1, 2, or 3 features, respectively. The regression weight associated with the number of features was greater than the corresponding regression weight in the Control condition, $t(34) = 2.12$, $p < .05$. In other words, the observation of exemplars that manifested the 75% feature base rates led participants to be more sensitive to similarity as compared to the Control condition, in which participants

were simply told about the 75% feature base rates.

Individual differences. On a finer level of analysis, there is some clustering in the data. Informally, the response patterns given by participants in the Causal Schema condition fell into a few different classes, and the most frequent was in fact the one that respects the causal Markov condition (uniformly low ratings when F_1 is absent, uniformly high ratings when F_1 is present). A look at the Causal Schema subjects' regression weights (see Figure 3) revealed a group of 8 "causal Markov" subjects who weighted the presence or absence of F_1 heavily and the number of features lightly, 3 "similarity" subjects who weighted the number of features heavily and F_1 lightly, and 7 "compromisers" who assigned moderate weights to both predictors. In contrast, examination of the Exemplars condition revealed 2 causal Markov subjects, 11 similarity subjects, and 2 compromisers (3 subjects weighted neither factor). That is, whereas the modal response in the Exemplars condition was similarity based, the modal response in the Causal Schema condition was consistent with the causal Markov condition.
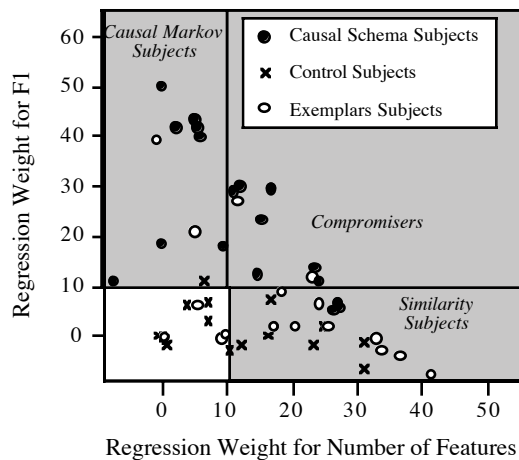

Figure 3. Regressions weights for individual subjects.

Categorization Results

One purpose of the categorization task was to ask to what extent feature inference was mediated by the goodness of category membership of the exemplar with the unobserved feature. We summarize the main findings. First, in the Causal Schema condition categorization ratings were influenced by the two factors that influenced feature predictions: They increased as the number of features in the exemplar increased (i.e., as the exemplar's similarity increased), and they increased even more when the causally central F_1 was present (i.e., the pattern shown for feature predictions in Figure 2a). However, unlike the prediction ratings, the categorization ratings were also sensitive to whether the causal relation between F_1 and each of the other observed features was confirmed or violated (i.e., whether F_1 and each of the effect features were jointly present/absent or not), a finding replicating past results (Rehder & Hastie, 2001). In other words, whereas categorization ratings showed a sensitivity to all three of the relations that constitute a common-cause schema, for purposes of predicting unobserved features participants apparently attended to only the relation in which

the unobserved feature dimension was involved. The result was a dissociation between feature prediction and categorization ratings in the Causal Schema condition.

Second, in the Exemplars condition category membership ratings, like the feature prediction ratings, were sensitive to similarity but insensitive to the presence or absence of F_1 as compared to other features (cf. Figure 2c). However, unlike the feature prediction ratings the categorization ratings were also sensitive to whether the correlations between F_1 and each of the other observed features were broken or preserved.

Importantly, this latter result speaks to the possibility that Exemplars participants' insensitivity to the presence or absence of F_1 during the feature prediction task arose merely as consequence of their failing to learn and encode the correlations involving F_1. In fact, results from the categorization task indicate that participants encoded these correlations but did not make use of them in feature inference. This represents a dissociation between feature inference and categorization, just as in the Causal Schema condition.

Finally, in the Control condition both category membership and feature prediction ratings were monotonic functions of the number of features present, that is, the featural similarity of the exemplar to the category prototype.

General Discussion

The first question asked in the current article was whether causal knowledge about a category influences predictions regarding the presence or absence of unobserved features. In fact, we found that causal knowledge had a strong effect on those inferences. Reasoners were much more likely to predict the presence of an unobserved feature when its cause was present than when that cause was absent. In this respect, their reasoning was similar to the normative method of inference defined by Bayesian networks.

These results contribute to a collection of findings demonstrating the importance of theoretical or explanatory knowledge in variety of feature inference tasks. For example, Lassaline (1996) found that the projection of a new property from one category to another was stronger when causal knowledge supportive of that property was provided (also see Sloman, 1994). Rehder & Ross (2001) found that the learning of a category via a feature prediction task proceeded more rapidly when features were related on the basis of prior knowledge. However, so far as we know, the current study is the first to address the specific role of causal knowledge in inferring the presence of unobserved features in a category with known causal structure.

Although causal knowledge had a profound effect on feature predictions, we also found that normative Bayesian reasoning is not the whole story. Even when reasoners had causal knowledge, their feature inferences showed a persistent effect of overall similarity to the category, such that an exemplar with a greater number of category-associated features was deemed more likely to have an unobserved feature. This was true even though the features that contributed to similarity were conditionally independent of the unobserved feature in question. In this respect, the effect of causal knowledge on feature predictions violated the causal Markov condition associated with Bayesian networks.

The influence of similarity on feature predictions also ob-

tained in the Control condition. This effect held even though participants did not observe category members (in contrast to previous studies demonstrating similarity effects, e.g., Yamauchi & Markman, 2000), but rather were provided only with a verbal statement of the 75% feature base rates. Under these conditions one might have expected that Control participants would be especially likely to assume independence among features (i.e., base each prediction only on the base rate of the feature in question and not on the presence/absence of other features). The current results indicate people's tendency to reason on the basis of central tendency or prototype information holds even when that information is provided in summary form (and without any mention of correlations between features) rather than experientially.

On the one hand, these findings are reminiscent of other studies that have attempted—mostly in vain—to induce participants to ignore the effects of similarity (e.g., Allen & Brooks, 1991). However, our analysis of individual differences revealed considerable variation among participants in the relative importance of causal knowledge and similarity. In fact, 8 of 18 participants in the Causal Schema condition ignored similarity (that is, they honored the causal Markov condition) when predicting unobserved features, indicating that similarity-based responding is not obligatory. The question of under what conditions feature inferences are dominated by theoretical and causal knowledge versus featural similarity is one that merits further investigation.

Another question we asked was whether causal knowledge has a different effect than inter-feature correlational knowledge. Rather than being given an explicit causal model, participants in the Exemplars condition observed exemplars that manifested the correlations implied by that model. In fact, though, the vast majority (14 of 18) of Exemplars participants failed to use those correlations in making feature inferences, and their inferences were qualitatively like those of Control subjects, who had neither causal nor empirical knowledge. Instead, the effect of the empirical observations was merely to make feature predictions even more sensitive to the degree to which an exemplar was similar to the category's prototype.

This finding contrasts with previous studies in which feature prediction was found to be sensitive to inter-feature correlations, at least when those correlations were observed during a classification-with-feedback task (Anderson & Fincham, 1996; Thomas, 1998; Yamauchi & Markman, 1998). We used a different learning task, in which participants were asked merely to observe category members, but we know that the inter-feature correlations were learned and encoded during this task because they were reflected in participants' categorization ratings. That the same participants failed to use these correlations in feature inference represents a dissociation between categorization and feature inference.

More generally, in both the Causal Schema and the Exemplars conditions, we found that categorization ratings and feature inferences were sensitive to different kinds of information: Categorization but not feature prediction was sensitive to the overall causal or correlational structure instantiated in an exemplar. This implies that feature inference is not merely mediated by goodness of category membership. Instead, participants in both of these conditions used the category knowledge they possessed in different ways depending on the task at hand. That is, whereas Yamauchi and Markman (1998) have suggested that category representations will differ depending on whether they are acquired via categorization or feature prediction, the current results suggest that categorization and feature prediction tasks can also draw on different aspects of a single representation.

References

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of the explicit rule. *Journal of Experimental Psychology: General, 120*, 3–19.

Anderson, J. R., & Fincham, J. M. (1996). Categorization and sensitivity to correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 259–277.

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

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference.* San Mateo, CA: Morgan Kaufman.

Rehder, B. (1999). A causal model theory of categorization. In *Proceedings of the 21st Annual Meeting of the Cognitive Science Society* (pp. 595–600). Vancouver, British Columbia.

Rehder, B. (2001). *A causal-model theory of conceptual representation and categorization.* Submitted for publication.

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General, 130*, 323–360.

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

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14*, 665–681.

Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology, 25*, 231–280.

Sloman, S. A. (1994). When explanations compete: The role of explanatory coherence on judgements of likelihood. *Cognition, 52*, 1–21.

Thomas, R. D. (1998). Learning correlations in categorization tasks using large, ill-defined categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 119–143.

Waldmann, M. R., Holyoak, K. J., & Fratianne, A. (1995). Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General, 124*, 181–206.

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

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

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

Yamauchi, T., & Markman, A. B. (2000). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 776–795.