Title
The Role of Coherence in Causal-Based Categorization

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 30(30)

ISSN
1069-7977

Authors
Rheder, Bob
Kim, ShinWoo

Publication Date
2008

Peer reviewed

The Role of Coherence in Causal-Based Categorization

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

Abstract

When features of categories are linked by causal relations, two different effects on classification judgments have been documented. The first, the *causal status effect*, is that features that appear earlier in a causal network (and thus are more "causal") are weighed more heavily in classification than less causal features. The second, the *coherence effect*, is that good category members are those whose combination of features make sense in light of causal laws (e.g., cause and effect feature both present or both absent). Recently, Marsh and Ahn (2006) suggested that previous studies (e.g., Rehder, 2003b) may have exaggerated coherence effects and minimized the causal status effect by use of the word "normal" to describe atypical feature values. We tested this hypothesis and found the opposite result, namely, avoiding the "normal" wording led to a *larger* coherence effect and a *smaller* causal status effect. Instead, coherence effects were extremely robust, never accounting for less than half of the variance in classification ratings induced by causal links.

A key aim of the study of concepts is to identify how facts we directly observe and theoretical beliefs involving explanatory and causal knowledge contribute to how we represent and use categories. Although early research into concepts focused on the effect of observations, subsequent research has shown that theoretical beliefs influence virtually every type of category-based judgment (see Murphy, 2002, for a review). This article is concerned with how one type of category-based judgment, classification, is affected by one type of theoretical knowledge, namely, causal knowledge that relates features of categories.

There are numerous examples of the causal relationships between category features. People know that having claws enables tigers to catch prey, having gills enables fish to breathe, having a fan allows an automobile's engine to remain cool. Numerous studies have investigated how this knowledge affects classification. Some studies have tested real-world categories (e.g., Ahn, 1998; Sloman, Love, & Ahn, 1998; Kim & Ahn, 2002), but in order to test alternative models, investigators have turned to artificial categories that are subject to experimental control. In these studies, participants are instructed on new types of objects and their features and causal relations among those features. They are then asked to judge the category membership of items displaying various combinations of features.

By testing different causal network topologies in this manner, researchers have uncovered several important effects of causal knowledge on categorization. We first describe two of those effects, namely, the *causal status effect* and the *coherence effect*. We then review recent evidence suggesting that unnatural stimulus materials in some previous studies might exaggerate the magnitude of one of these effects (the coherence effect) at the expense of the other (the causal status effect). We then present the results of an experimental test of this claim.

Causal Status Effect

The causal status effect is the phenomenon in which, all else being equal, features that appear earlier in a category's causal network (and thus are "more causal") carry greater weight in categorization decisions. For example, in Fig. 1 F_1 is the most causal feature, F_4 is the least causal, and F_2 and F_3 are intermediate. As a consequence, F_1 should be weighed more than F_2, which should be weighed more than F_3, which should be weighed more than F_4. Of course, the causal status effect does not imply that features' categorization weight is only determined by their causal status, because it is well known that those weights are also influenced by features' salience (Sloman et al. 1998) and their *cue validity* (the extent to which they are diagnostic of that category versus another, Rosch & Mervis, 1975). But the claim is that causal status will dominate when these factors are equated.

One study that provides partial support for the causal status effect was conducted by Rehder (2003b). Participants were instructed on, for example, Myastars (a type of star) that had four features causally related in causal chain. Participants were then shown a series of stars and asked to rate (on a 0-100 scale) if they were a Myastar. Rehder performed regression analyses to assess the influence that each feature had on those ratings. The left panel of Fig. 2A presents the feature regression weights from that study. Consistent with a causal status effect, feature F_1 was weighed more heavily than feature F_2. F_1's weight of 9.6 meant that categorization ratings were about 10 points higher (all else being equal), when a test item possessed F_1 and 10 points lower when it did not. In contrast, F_2's weight of 6.2 meant that ratings experienced a swing of only about 12 points depending on whether F_2 was present or absent. But whereas F_1 was weighed more than F_2, the weights of features F_2, F_3, and F_4 were virtually identical, indicating that only a partial causal status effect obtained. This result contrasted with other studies reporting a full causal status effect (e.g., with a three-feature chain network in Ahn et al. 2000).

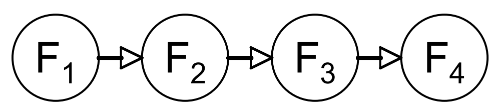

Figure 1

In considering reasons for this partial causal status effect, Rehder noted that participants were given very concrete in-

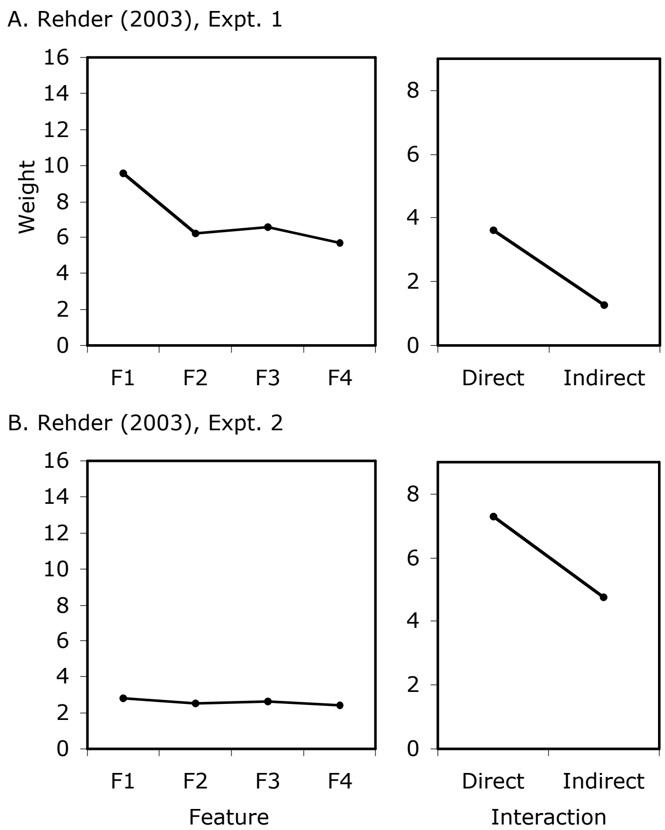

Figure 2

formation about features' base rates within the categories. For example, participants were told that "75% of Myastars have high temperature whereas 25% have a normal temperature," "75% of Myastars have high density whereas 25% have a normal density," and so on. A partial causal status effect may have obtained because subjects were reluctant to weigh the features differently given the numerically identical base rates. To test this possibility, a second experiment omitted base rate information (e.g., "Some Myastars have a high density whereas others have normal density."). The results presented in the left panel of Fig. 2B indicate that feature weights were close to 0, an expected result given that participants were given no reason to think that a star with high density was more likely to be a Myastar than one with normal density. More importantly, there was no sign of a causal status effect, as features were weighed equally. These results indicate that the 75% base rate information provided in Expt. 1 was not responsible for its lack of a full causal status effect. Later we will discuss other variables that may have an effect on the causal status effect, but we first turn to the second phenomenon, the coherence effect.

Coherence Effect

Because Rehder (2003b) asked participants to rate all possible exemplars that could be formed on the category's binary dimensions, he was able to assess not just feature weights but also feature *interactions*, that is, how important certain *combinations* of features are to category membership. This was done by introducing interaction terms in the regression equation for each pair of features (i.e., F_1F_2, F_2F_3, F_3F_4, F_1F_3, F_2F_4, and F_1F_4). For example, the F_1F_2 interaction term codes whether features F_1 and F_2 are both present or both absent versus one present and the other absent. The regression weight on that interaction term represents the importance to participants' categorization ratings of dimensions F_1 and F_2 having the same value (present or absent) or not.

The right-hand panels of Fig. 2 indicate that in both Expts. 1 and 2 participants were quite sensitive to whether potential category members exhibit not just the right features considered individually but also the right combination of features. In Fig. 2 interactions have been grouped into two types, the *direct* interactions between features that are directly causally related (F_1F_2, F_2F_3, and F_3F_4) and the *indirect* interactions between the features that are only indirectly related (F_1F_3, F_2F_4, and F_1F_4). Fig. 2 shows that both types of interaction weights were greater than zero. The direct interaction weight of 7.3 in Expt. 2 indicates, for example, that categorization ratings were about 7 points higher (all else being equal), when a test item possessed either both F_1 and F_2 or neither one, and 7 points lower when it possessed one but not the other. That is, participants were sensitive to the interfeature correlations one would expect in light of the causal relations, so that an item was more coherent and thus a better category member if it maintained the expected correlations between F_1 and F_2 (and F_2 and F_3, and F_3 and F_4). This interpretation is reinforced by the finding of lower but still positive weights on the indirect interaction terms, because one expects variables that are indirectly causally related to be correlated, albeit not as strongly as directly-linked variables (at least for probabilistic causal links). Apparently, good category members are those that manifest the complex pattern of correlations one expects to be generated by a causal network. This sensitivity to coherence has been found in numerous studies (Marsh & Ahn, 2006), including those testing other causal networks (Rehder & Hastie, 2001; Rehder 2003a; Rehder & Kim, 2006).

Another notable result in Fig. 2 is that the coherence effect was larger in Expt. 2 than in Expt. 1. A straightforward interpretation of this result is that, because participants could give little weight to individual features in Expt. 2 (because of the absence of feature base rate information), coherence became the only factor available to determine their categorization rating

Questions About the Coherence Effect

Despite the considerable evidence in favor of the coherence effect, questions about its robustness have been raised. Recently, Marsh and Ahn (2006) offered a critique of Rehder (2003b) in which they questioned both the presence of large coherence effects and the absence of a full causal status effect, arguing those findings may have been due to an artifact of the experimental materials. As mentioned, in Rehder (2003b) participants were told that Myastars could have either high or normal temperature, high or normal density, etc. Marsh and Ahn argued that the use of "normal" might be problematic, for two reasons. First, although the intent of the "normal" wording was to define Myastars with respect to all stars, Marsh and Ahn argued that the use of "normal" might have inflated coherence effects because participants expected the "normal" values to appear together.

Second, Marsh and Ahn noted that Rehder's Expt. 2 pro-

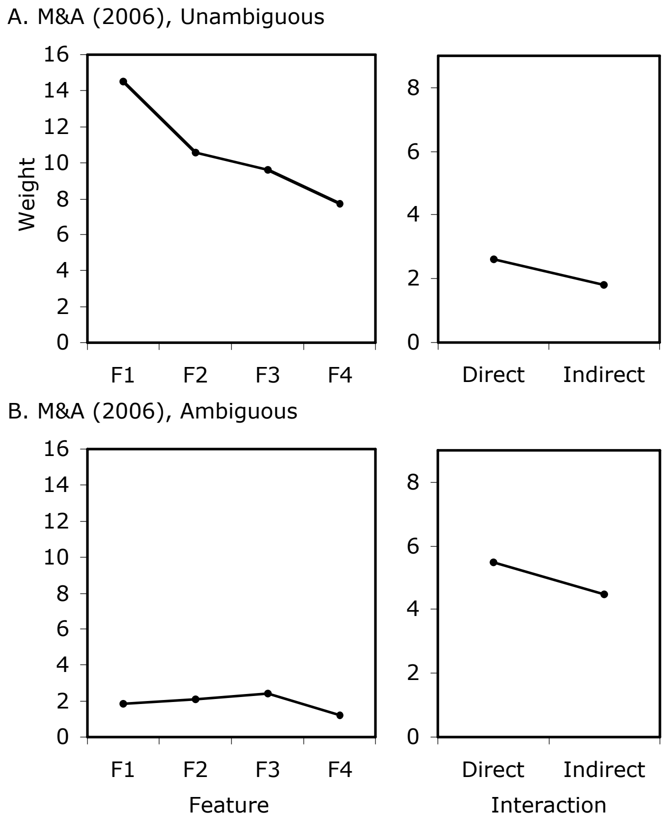

Figure 3

vided no information about feature base rates. In the absence of base rate information, items that had all "normal" values (i.e., the "0000" item, where 0 = an uncharacteristic value on a binary dimension) were given high categorization ratings. They argued that 0000's high rating was manifested in the form of unusually high two-way interaction terms, that is, a large coherence effect. (Of course, this second point applies to Rehder's Expt. 2 but not Expt. 1, in which base rate information was provided.)

To test these hypotheses, Marsh and Ahn tested participants in what they referred to as the Ambiguous and Unambiguous conditions. The Ambiguous condition was intended to be a replication of Rehder (2003b), Expt. 2 ("ambiguous" because of the ambiguous base rates). The Unambiguous condition differed from the Ambiguous condition in two ways. First, participants were given feature base rate information (e.g., "Myastars tend to have high density."). Second, the controversial "normal" wording was avoided. For example, rather than to-be-rated stars having either a hot temperature or a normal temperature, they had a hot temperature or a *low* temperature; rather than high density or a normal density, they had high density or *low* density; and so on. Participants then performed a classification rating task essentially identical to the one in Rehder (2003b).

The regression weights from the Unambiguous and Ambiguous conditions are presented in Fig. 3. Fig. 3B indicates that, despite some differences in the experimental procedure (e.g., participants were tested on multiple categories, a diagram on the causal relations was presented on each classification test trial, etc.), the Ambiguous condition essentially replicated the findings of Rehder (2003b), Expt. 2 (Fig. 2B). Results in the Unambiguous condition, in contrast, showed a markedly different pattern. First, the individual weights were much larger and exhibited an almost full causal status effect (F_1 > [F_2, F_3] > F_4). And, although both interaction weights were positive, the effect of coherence was lower than in the Ambiguous condition.

Marsh and Ahn interpreted these results as supporting their claim that the absence of base rate information and the use of the "normal" attribute values were responsible for the equivocal causal status effect and strong coherence effect found in Rehder (2003b). They concluded that "…[Rehder's] results are methodological artifacts arising from the unnatural wording of category attributes" [p. 561].

Analysis of Marsh & Ahn (2006)

The possibility that reported effects of causal knowledge on classification might be due to specific stimulus materials is of course is an important one, and in this light Marsh and Ahn's (2006) study is a welcome contribution to the literature. But do their empirical results warrant their negative conclusion regarding Rehder (2003b)? To evaluate this claim, it is important to remember that they tested conditions that differed in two ways: as compared to the Ambiguous condition, the Unambiguous condition provided base rate information *and* avoided use of the "normal" wording. Thus, to understand the theoretical significance of their study, we must assess to which of these manipulations the empirical findings should be attributed.

Regarding the base rate manipulation, we believe there is no reason to doubt that people's classification performance strongly depends on their beliefs about features' base rates. Indeed, this has been a standard finding in the categorization literature for decades (Rosch & Mervis, 1985). Moreover, the effect of base rates in the presence of causal knowledge is already part of the literature; in fact, it is part of the Rehder (2003b) study that Marsh and Ahn aimed to criticize. Recall that whereas in Rehder's Expt. 1 participants were told that typical features occurred in 75% of category members, no base information was provided in Expt. 2. A comparison of Figs. 2A and 3A indicates that the effect of this base rate information was almost exactly the same as that seen by Marsh and Ahn. On the importance of feature base rates then, there is no controversy.

However, Marsh and Ahn reached an additional conclusion, namely, that the differences between their conditions were also due to the "normal" wording. Although this conclusion might seem puzzling given that the results of Rehder's Expt. 1 (Fig. 2A) that used the "normal" wording were so similar to their Unambiguous condition (Fig. 3A) which did not, note that whereas Rehder's Expt. 1 found that F_1 > [F_2, F_3, F_4], in the Unambiguous condition it was also the case [F_2, F_3] > F_4; that is, a more complete causal status effect obtained. For Marsh and Ahn, this difference was sufficient to conclude that the use of "normal" can suppress the causal status effect: "We argue that [the results from Rehder's Expt. 1] occurred because even with the high base-rate information the non-prototype values were still described as "normal"…highlighting the violations of the correlated structure to the participants." (p. 566).

Overview of Experiment

Is there an effect of using "normal" feature values and does it influence either the causal status or coherence effects? Despite Marsh and Ahn's (2006) claim, it is impossible to say on the basis of their study. First, comparing their Unambiguous and Ambiguous conditions is of no help, because, as noted, those conditions also differed on the base rate information provided. And comparing their Unambiguous condition with Rehder's Expt. 1 is problematic because those studies also had numerous procedural differences (within- vs. between subject design, base rates of "75%" vs. "tend to have," a diagram was provided on the classification test trial in one study but not the other, etc.).

To test what effect, if any, "normal" feature values have on causal-based classification, we compared a Normal condition in which the non-typical values were described as "normal" (as in Rehder, 2003b) and a Bipolar condition in they were described as opposite to the typical value (as in Marsh and Ahn's Unambiguous condition). The conditions were otherwise identical. Feature base rate information was provided in both conditions.

Method

Materials. The six novel categories tested were the same as those in Rehder (2003b) and Marsh and Ahn (2006). For example, the features of Myastars were *made of ionized helium*, *hot temperature*, *high density*, and *large number of planets*. The base rate for each typical feature was described as "most" and the non-typical value was described as "normal" in the Normal condition and as a value opposite to the typical value in the Bipolar condition. For example, participants were told "Most Myastars have high density whereas some have a normal density." in the Normal condition and "Most Myastars have high density whereas some have low density." in the Bipolar condition. All participants learned three causal link arranged as in Fig. 1. Each causal relationship specified the cause and effect features (e.g., "Ionized helium causes the star to be very hot.") and some detail regarding the causal mechanism (e.g., "Ionized helium participates in nuclear reactions that release more energy than the nuclear reactions of normal hydrogen-based stars, and the star is hotter as a result.").

Procedure. The procedure was the same as in Rehder (2003b). Participants first studied several screens of information about their assigned category at their own pace. They were then required to pass a multiple-choice test to ensure they learned the information. Participants then rated on a 0–100 scale the category membership of all possible 16 objects that can be formed on four binary dimensions. On each trial the four features were listed in dimension order (1–4) on the computer screen. The order of the trials was randomized for each participant.

Participants. 72 New York University undergraduates received course credit for participating in this experiment. They were assigned in equal numbers to the Normal and Bipolar conditions and to one of the six categories.

Results

Following Rehder (2003b) and Marsh and Ahn (2006) category membership ratings were analyzed by performing a multiple regression for each participant. The regression weights averaged over participants in each condition are presented in Fig. 4.

First consider the feature weights shown in the left panels of Fig. 4. In the Normal condition, those weights were large and exhibit a partial causal status effect in which feature F_1 was weighed more heavily than the other features which in turn did not differ. These results are qualitatively identical to those found in Rehder (2003b), Expt. 1, which also used "normal" as the atypical feature values (Fig. 2A).

The question of interest of course is how these results compare with the condition in which "normal" was not used, that is, in the Bipolar condition presented in Fig. 4B. In fact, the figure indicates the complete *absence* of a causal status effect in the Bipolar condition. That is, rather than enhancing the causal status effect, use of an opposite value as the atypical value (*low* vs. high density) eliminated any tendency for participants to weigh more causal features more heavily than less causal features. This of course is exactly the *opposite* of the Marsh and Ahn (2006) claim.

The magnitude of the coherence effect in these conditions (right panels of Fig. 4) tells a similar story. In the Normal condition, both the direct and indirect interaction terms are positive and about the same magnitude as in Rehder's Expt. 1 (albeit the direct term is a bit lower and the indirect term a bit higher). In contrast, the interaction terms are over twice as large in the Bipolar condition, indicating a much larger coherence effect in that condition. That is, rather than the "normal" wording exaggerating coherence effects in fact it *minimizes* that effect. Again, this effect is precisely the reverse of what Marsh and Ahn argued.

These conclusions were supported by statistical analysis. A 2 x 4 mixed ANOVA of the feature weights was conducted where the between-subject factor was condition (Normal vs. Bipolar) and the within-subject factor was feature (1–4). There was a main effect of feature, $F(3, 210) = 8.20$, $MSE = 21.8$, $p < .0001$, reflecting different feature weights, and an interaction between features and condition, $F(3, 210) = 8.20$, $MSE = 21.8$, $p < .0001$, indicating that the pattern of weights varied across condition. A separate analysis of the Normal condition indicated that feature F_1 was weighed more heavily than the other three features, $F(1, 35) = 21.65$, $MSE = 21.8$, $p < .0001$, which in turn did not differ from one another. In the Bipolar condition, feature weights were not significantly different. In addition, a 2 x 2 ANOVA of the interaction weights was conducted with condition (Normal vs. Bipolar) and interaction term (direct vs. indirect) as factors. There was a main effect of condition, $F(1, 70) = 13.63$, $MSE = 41.1$, $p < .001$, reflecting the larger interaction weights in the Bipolar condition. There was also a main effect of term, $F(1, 70) = 6.46$, $MSE = 6.6$, $p < .01$, indicating that the direct terms were larger than the indirect ones. The interaction was not significant.

Discussion

This experiment assessed whether use of "normal" for atypical feature values affects causal-based classification. The answer is that it does. As compared to bipolar dimen-

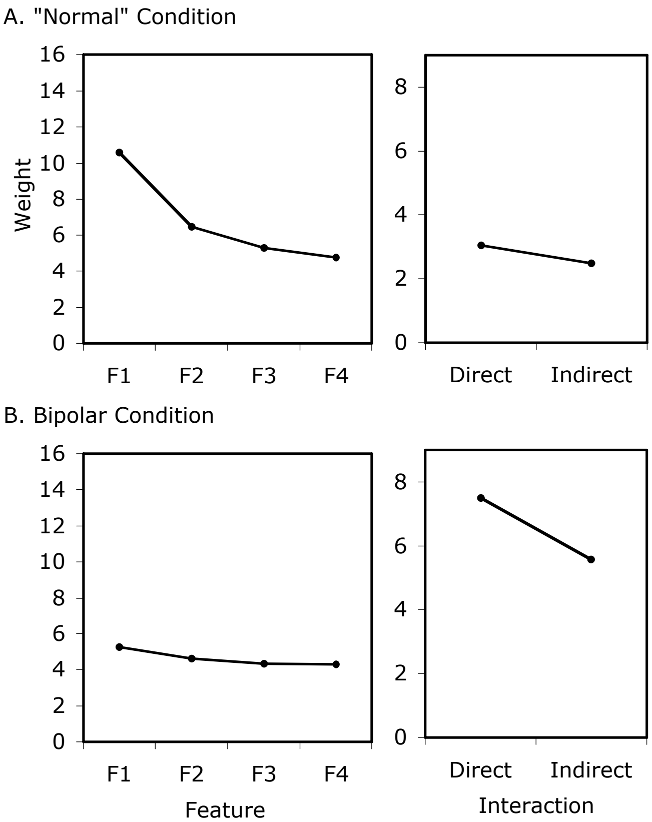

Figure 4

sions, use of "normal" enhances the causal status effect and decreases the coherence effect. This of course is just the reverse of Marsh and Ahn's (2006) claim regarding the effect of the "normal" values.

Why might bipolar dimensions lead to stronger coherence effects? One possibility is that such dimensions might lead participants to infer the existence of additional causal links. For example, if you are told that Myastars have either high or low temperature and either high or low density, and that high temperature causes high density, you might take this to mean that *low* temperature also causes *low* density. On this account, a star with high temperature and low density violates two causal links rather than one. Or, you might treat temperature and density as continuous variables, in which case high temperature and low density is a more egregious violation of the causal link than high temperature and *normal* density (which implies that density is at the midpoint of the density scale). But whatever the reason, it is clear that bipolar dimensions highlight rather than diminish the importance of the correlational structure among features.

The absence of a causal status effect in the Bipolar condition might seem puzzling in light of the results in Marsh and Ahn's Unambiguous condition which also tested bipolar dimensions. Recall, however, that there were numerous procedural differences between experiments: base rates conveyed as "tend to have" vs. "most," within- vs. between-subjects designs, one study presented a diagram during the classification trial, etc. Apparently, one or more of these differences must have some effect on the causal status effect. The important point however is that when these sorts of procedural variables are equated, the effect of bipolar dimensions is to enhance the coherence effect and minimize the causal status effect.

Note that Marsh and Ahn also tested two other causal networks, namely, a common cause network (one feature causes three others) and common effect network (one feature caused by three others) and found results analogous to those with a chain network: low feature weights and high interaction terms in the Ambiguous condition (replicating Rehder, 2003a) and higher feature weights and lower interaction terms in the Unambiguous condition. But their conclusion that these results were partly due to the "normal" wording ignores the fact that those conditions also differed on the base rate information provided. Now that we know that, all else being equal, "normal" *increases* coherence effects, it must be that the Marsh and Ahn's Unambiguous condition's lower coherence effect was solely due to its base rate information. Other differences between their Unambiguous and our lab's own common cause and common effect conditions that provided base rate information (e.g., in Rehder & Hastie, 2001) are likely due to the procedure differences that we have mentioned.

General Discussion

The purpose of this research was to evaluate the claim that previous findings regarding the causal status and coherence effects in causal-based categorization were invalid due to "methodological artifacts," namely the use of the word normal to describe atypical dimension values. According to Marsh and Ahn (2006), use of "normal" was likely to have exaggerated the importance of coherence and minimized the importance of features' causal status. However, this conclusion was based on comparing conditions that differed in ways other than the "normal" wording. We tested conditions that differed only in the "normal" wording and found just the opposite effect. That is, the use of "normal" *minimizes* coherence effects and *enhances* the causal status effect, not the other way around.

It is important to remember that Marsh and Ahn made two claims, and with the second of those claims—that coherence effects will be larger when feature base rate information is omitted—we have no disagreement. It is clear that if coherence is the only basis on which to make a categorization judgment, then participants will use most of the response scale to express how test items differ in their coherence and, as a result, the magnitude of the interaction terms will be larger. We also agree that people usually have some idea of what features are typical of a category. But recall that Rehder (2003b, Expt. 2) omitted base rate information as a follow-on to a first experiment to see if a full causal status effect would emerge. Another study from our lab omitted base rate information (Rehder, 2003a) to make a detailed study of the feature interactions produced by more complex causal networks. But of course real-world categories generally exhibit a family resemblance structure and to mimic this structure all other studies in this line provided base rate information (Rehder, 2003b, Expt. 1, Rehder & Hastie, 2001; Rehder & Kim, 2006). Despite the presence of this base rate information, coherence effects were found in all of them. Of course, the present results suggest that had bipolar dimensions been used instead of "normal," the coherence effects in those studies would have been larger still.

In summary then, there is no reason to think that the coherence effect is a "methodological artifact" of stimulus materials. This finding is important because coherence has been shown to affect every category-based judgment in which it has been assessed. Coherence affects category-based induction, with more coherent items supporting stronger inductions (e.g., Patalano & Ross, 2007; Rehder & Hastie, 2004). It affects learning, with more coherent categories being easier to learn (e.g., Murphy & Allopenna, 1984). Coherence also affects categorization even when feature relations are neither explicitly causal nor provided as part of the experiment (Wisniewski, 1995). In other words, coherence is one of the most important manifestations of knowledge-based (or "theory-based") categories.

Moreover, within causal-based classification coherence can be shown to dominate other effects such as changes to individual feature weights. To demonstrate this, we computed the proportion of the variance in categorization ratings attributed to it and the causal status effect. The total variance induced by causal knowledge was taken to be the variance explained by a regression model with separate predictors for each feature and each two-way interaction as compared to a model with only one predictor representing the total number of characteristic features. The variance attributable to the causal status effect is the variance explained by the separate predictors for each feature in the full model whereas that attributable to the coherence effect is the variance explained by the interaction terms. The results of this analysis are presented in Table 1. Even in our conditions in which a causal status effect was present, the coherence effect always accounted for more than 80% of the variance. When no causal status effect was present, coherence accounted for over 99% of the variance. And despite Marsh and Ahn's (2006) claim that "individual features' causal status, rather than feature combinations, was the predominant determinant" of categorization performance in their Unambiguous condition (p. 566), coherence accounted for 56% of the variance in that condition as well. Indeed, their suggestion that "a strong case for the role of inter-feature links has yet to be made" (p. 566) is especially puzzling given that such a "strong case" was present in their own data. Instead, Table 1 indicates that in these sorts of experimental paradigms the most important factor influencing categorization judgments is whether an object displays a configuration of features that make sense in light of the category's causal laws. Compared to coherence, changes to the importance of individual features (e.g., the causal status effect) run second.

Table 1

	% Variance Explained	
Study	Causal Status Effect	Coherence Effect
Rehder (2003b)		
Experiment 1	17.8%	82.2%
Experiment 2	0.9%	99.1%
Marsh & Ahn (2006)		
Unambiguous	43.6%	56.4%
Ambiguous	0.3%	99.7%
This study		
Normal	30.1%	69.9%
Bipolar	0.3%`	99.7%

References

Ahn, W. (1998). Why are different features central for natural kinds and artifacts? The role of causal status in determining feature centrality. *Cognition, 69*, 135-178.

Ahn, W., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology, 41*, 361-416.

Kim, N. S., & Ahn, N. S. (2002). Clinical psychologists' theory-based representation of mental disorders affect their diagnostic reasoning and memory. *Journal of Experiemental Psychology: General, 131*, 451-476.

Marsh, J., & Ahn, W. (2006). The role of causal status versus inter-feature links in feature weighting. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 561-566).

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

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

Patalano, A. L., & Ross, B. H. (2007). The role of category coherence in experience-based prediction. *Psychonomic Bulletin & Review, 14*, 629-634.

Rehder, B. (2003a). Categorization as causal reasoning. *Cognitive Science, 27*, 709-748.

Rehder, B. (2003b). A causal-model theory of conceptual representation and categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1141-1159.

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., & Hastie, R. (2004). Category coherence and category-based property induction. *Cognition, 91*, 113-153.

Rehder, B. & Kim, S. (2006). How causal knowledge affects classification: A generative theory of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 659-683.

Rehder, B. & Kim, S. (2008). *Causal status and coherence in causal-based categorization*. Submitted for publication.

Rosch, E. H., & Mervis, C. B. (1975). Family resemblance: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573-605.

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

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