

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Causal Status and Explanatory Goodness in Categorization

Permalink

<https://escholarship.org/uc/item/01f925kc>

Journal

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

ISSN

1069-7977

Authors

Jameson, Jason
Gentner, Dedre

Publication Date

2008

Peer reviewed

Causal Status and Explanatory Goodness in Categorization

Jason Jameson & Dedre Gentner

Northwestern University
Department of Psychology, 2029 Sheridan Road
Evanston, IL 60208 USA

Abstract

Much research (e.g., Keil, 1989; Murphy & Medin, 1985; Rips, 1989) has emphasized the critical role that domain knowledge plays in categorization judgments. Recent instantiations of this view (e.g., Ahn, et al., 2000; Rehder & Hastie, 2001) have focused on characterizing how *causal knowledge* supports categorization decisions. We suggest that a more satisfactory account of categorization can be gained by considering the broader role that causal information plays in processes of *explanation*. An explanation-based perspective treats categorization as an inference to the best explanation (Murphy & Medin, 1985; Rips, 1989). This suggests that a critical source of constraints on categorization may come from a direct investigation of *explanatory goodness*. We present evidence that the causal status effect (e.g., Ahn, et al., 2000)—i.e., the phenomenon in which causes tend to be more heavily weighted than effects in categorization judgments—depends on the goodness of the explanation in which the causal information is embedded.

Keywords: Categorization, causal status, explanation, explanatory goodness.

Introduction

Categorization processes are fundamental in making sense of the world. Not surprisingly, the study of categorization is of central importance in cognitive science; researchers have long sought to characterize the nature of the conceptual structures that best support classification decisions. There is general agreement that some parts of a concept are more important—more *conceptually central*—than others for categorization (e.g., Ahn, 1998; Ahn, et al., 2000; Medin & Shoben, 1988; Rehder & Hastie, 2001). For example, as Medin and Shoben (1988) demonstrated, the property of “curvedness” is more central for the category of boomerangs than for the category of bananas. That is, “straight bananas” are better examples of bananas than “straight boomerangs” are of boomerangs, because “curvedness” plays a more central role in the theoretical principles supporting one’s knowledge of boomerangs.

We argue here that conceptual centrality is closely related to the goodness of an explanation: specifically, that the centrality of a conceptual component derives from its role in the best explanation of what holds category members together. On this view, a good conceptual structure is one that contains important explanatory information.

The study of explanation spans several disciplines, ranging from philosophy and computer science (e.g., Thagard, 1989) to psychology (Read & Marcus-Newhall; Lombrozo, 2006; Lombrozo & Carey, 2006; Rips, 1989). In

psychology, the view that category knowledge comprises important explanatory information is not new; and our proposal takes off from the theory- or knowledge-based framework of conceptual structure (e.g., Carey, 1985; Keil, 1989; Murphy & Medin, 1985). We also draw on a specific instantiation of this view, the causal status hypothesis (Ahn, 1998; Ahn, et al., 2000). Murphy and Medin (1985) argued persuasively that categorization can be viewed as an inference to the best explanation (see also Rips, 1989). On the relationship between theoretical knowledge and centrality, Murphy and Medin defined the *weight* associated with features as “[d]etermined in part by importance in the underlying principles” (p. 298, 1985). This characterization is best viewed as a starting point, for it leaves open the critical question of how to define “importance.” One prominent answer to this question is the *causal status hypothesis*, according to which (1) the role that a feature plays in a concept—either as a cause or as an effect—partly determines its centrality; and (2) causal features are more important than effect features for categorization judgments: e.g., a potential category members’ possession of causal features is more diagnostic of category membership than possession of effect features. This hypothesis has received support in a variety of contexts and for several cognitive tasks (see Ahn & Kim, 2000, for a review).

Our central claim is that the causal status effect derives from the supporting role that causal information plays in explanation. Although this idea is a long-standing theme in the study of categorization (e.g., Murphy & Medin, 1985; Rips, 1989), relatively little research has directly examined how explanatory goodness relates to categorization. In this paper, we study the relationship between explanatory goodness and categorization. We demonstrate that the magnitude of the causal status effect depends on the quality of the explanation in which the causal information is embedded.

We first briefly review the causal status hypothesis and present arguments that the causal status effect depends on explanatory goodness. Then we present a study to support this claim. Finally, we discuss the implications of this study for categorization more broadly.

Conceptual centrality

Intuitively, some components of a concept do indeed appear more important for categorization judgments. As Medin and Shoben (1988) point out, it is much easier to imagine a robin that is not red than it is to imagine a robin that lacks the appropriate genetic structure of robins:

imagining a robin with, say, zebra DNA, would make many other typical properties of robins implausible (Wilson & Keil, 1998).

There is a long history of research on categorization demonstrating that some features associated with category members are particularly important for categorization judgments (e.g., Ahn, et al., 2000; Medin & Shoben, 1988; Rosch & Mervis, 1975; Tversky, 1977). One prominent perspective on centrality is based on the theory- or knowledge-based view of conceptual structure (e.g., Carey, 1985; Keil, 1989; Murphy & Medin, 1985; Rips, 1989). According to this framework, knowledge of a category is, in important ways, like a scientific theory, comprising a “host of mental explanations” (Murphy & Medin, 1985). In this view, category knowledge is seen, not in terms of a prototypical member or collection of exemplars, but rather in terms of “an explanatory principle common to category members (p. 298, Murphy & Medin, 1985). Accordingly, categorization is typically viewed as an inference to the best explanation. Murphy and Medin illustrate this idea with a well-known example: a man at a party who jumps into a pool fully clothed in a business suit would probably be classified as intoxicated, not because he is similar to the prototype or to instances of a “drunken behavior” category, but because being intoxicated is the best explanation for his behavior (see also Rips, 1989, for a clearly articulated sketch of this idea).

Despite the importance of the theory-based view for orienting research on categorization and centrality, the framework leaves unspecified the specific constraints related to background knowledge that influence classification decisions. One important response to this challenge is the idea that information that participates in causal relations is of greatest centrality (Ahn, et al., 2000; Rehder & Hastie, 2001; Rehder & Burnett, 2004). In Ahn’s (1998) *causal status hypothesis*, causes are weighted more heavily in categorization decisions than are the corresponding effects. For example, in the simple case of a single causal relationship between two features, such as having wings and flying (with wings enabling flying), having wings would be given greater weight than flying, and would more greatly influence categorization decisions.

Why should causes play a more important role than the corresponding effects in the underlying principles of a category? According to Ahn, et al. (2000), causal properties may be seen as generating other features: e.g., DNA produces external features such as hair color. Features such as DNA may thus be regarded psychologically as most defining or diagnostic of category membership, because they form part of the essence or core of a concept. It follows, Ahn, et al. claim, that possessing the most central features would provide better evidence for category membership than more superficial features.

While not disputing the empirical validity of the causal status effect, and the critical role of causal information for many if not most knowledge-dependent categorization decisions, we worry that the exclusive focus on causal

knowledge risks obscuring the supporting role that causal information plays in explanatory processes. As Lombrozo (2006) notes, explanations can’t be reduced to just the supporting causal information, as explanations entail a set of factors that go beyond the causal information per se (see also Read & Marcus-Newhall, 1993). For one thing, there are many kinds of causal structures (Forbus & Gentner, 1986; Rottmann & Gentner, 2006)—e.g., causal chains versus feedback systems—that draw on different kinds of domain knowledge. An additional factor is the preference for simple or parsimonious causal explanations over those that would invoke more assumptions (Lombrozo & Carey, 2006; Read & Marcus-Newhall, 1993; Rips, 1989).

In a recent study that illuminates the relationship between the causal status effect and explanatory structure, Lombrozo (2007b) has shown that the strength of the causal status effect depends partly on the type of explanation—mechanistic (an appeal to direct, proximate causes) or teleological (an appeal to function)—in which the causal knowledge is embedded. In one experiment, participants were presented with a novel category characterized by two features and were told that one feature (e.g., eats blueberries) causes the other (e.g., has blue fur). Participants were then asked to explain why the animal has the second feature (i.e., blue fur), and to estimate the probability that an object missing one of the features was a member of the category. Participants who provided a mechanistic explanation (e.g., mentioning eating blueberries as the proximate cause) showed a larger causal status effect than those who provided a teleological explanation (e.g., blue fur serves as camouflage).

In addition to the *type* of explanation, the causal status effect may also depend on the *quality* of an explanation. We suggest that some of the puzzling results obtained in studies of the causal status hypothesis might be best explained in terms of explanatory goodness. In one such study (Experiment 4), Ahn, Kim, Lassaline, and Dennis (2000) gave participants a standard (a single sentence describing a causal relationship between a cause and two properties treated as effects) and two alternatives. Participants were instructed to choose which of the two alternatives should be categorized with the standard. One standard read: “This object has a high-intensity light bulb and a pouch that can contain liquid because it was designed to kill bugs.” The first alternative shared the cause, but not the effects (i.e., “This object has a sweet, smelly patch and an x-ray generator because it was designed to kill bugs.”); the other alternative shared the effects, but not the cause (i.e., “This object has a high-intensity light bulb and a pouch that can contain liquid because it was designed to be used in a photograph studio.”).

Ahn, et al. predicted that people would prefer to match an alternative to the standard on the basis of the shared cause, rather than the shared effects. Although the overall results confirmed the prediction, inspection of the results reveals that only half the items conformed to the prediction. Table 1

shows the materials used in the study and, for each item, the percentage of choices for the matching cause.

Table 1: Experiment 4 standards, from Ahn, et al., (2000).

Standard	%
This object has a high-intensity light bulb and a pouch that can contain liquid because it was designed to kill bugs.	80
This sculpture is made of metal and consists of six cubes stacked up because the sculptor intended it to symbolize pollution.	72
This object has a rubber platform and vibrates smoothly because it was designed to relax pregnant mares during labor.	60
This plant has needle leaves, and produces tiny pink flowers in the spring because it has a DNA structure called Valva.	52
This animal has a block-shaped head, is red, and has 13 teeth because this animal has a genetic code, XB12.	52
This painting has four pillars and is red because the painter intended to draw a dog.	40

While there is a noticeable causal status effect for three of the items, two items exhibit chance responding, and one even suggests a preference for the shared effect alternative.

We suggest that the item variability can best be understood in terms of the explanatory role of the causal property being asserted for each item. One hallmark of a good functional explanation is that the intended function is causally connected to the effects to be explained. Ideally, the facts can then be understood as subsumed under a general causal law (Lombrozo & Carey, 2006). In addition, good explanations exhibit breadth—the extent to which an explanation accounts for most if not all of the available facts (Pennington & Hastie, 1988; Read & Marcus-Newhall, 1993; Thagard, 1989); and depth—the extent to which the local explanation fits within a larger explanatory framework, and can itself be explained (Read & Marcus-Newhall, 1993; Thagard, 1989).

These considerations can shed light on why some items in Ahn et al.’s study conformed to their prediction, while others did not. Consider the item with the strongest causal status effect, the first item (“This object has a high-intensity light bulb and a pouch that can contain liquid because it was designed to kill bugs.”). In this case, a clear, general causal relationship links the intention with the effects: it’s generally known that bugs are attracted to light, and that

bug-killing agents typically take the form of liquid pesticides. Second, the explanation is broad (intention to kill bugs explains every stated effect) and deep (the intention to kill bugs fits within a larger explanatory framework: bugs are generally annoying).

Now, consider the worst-performing item: “This painting has four pillars and is red because the painter intended to draw a dog.” The causal relation between four pillars and the intention to draw a dog is extremely weak, and there is no clear explanatory framework that would link the intention to paint a dog with any of the facts.

Of course, these are after-the-fact suggestions. To make this account plausible, what is needed is a manipulation of explanatory goodness. In the current study, we test the hypothesis that the magnitude of the causal status effect depends on explanatory quality: we predict that the better the explanation in which causal information is embedded, the stronger the causal status effect.

Experiment

To test the hypothesis that the magnitude of the causal status effect depends on explanatory quality, we manipulated the quality of explanations in a two-alternative forced choice categorization task, as in Ahn, et al.’s (2000) study. Thus, participants had to choose which of two alternatives should be categorized with a given standard. As a manipulation check, we also collected judgments of explanatory quality for fact-explanation pairs: pairs of sentences that describe facts and a potential explanation for those facts. The sentences were simple and clearly interpretable, patterned after the materials used in Ahn, et al., (2000).

Method

Participants Twenty undergraduate students from Northwestern University participated for partial credit in an introductory psychology course.

Design, materials, and procedure The design was a 2×2 mixed factorial with explanatory quality (good or poor), manipulated within-subjects and task order (categorization task first or second), manipulated as a between-subjects counterbalancing variable. The materials consisted of 24 stimulus sets, each based on a standard “fact-explanation” pair, which comprised a set of facts joined with a potential explanation for those facts. In the explanation-rating task the facts were presented as short sentences describing a particular situation, which participants were instructed to assume as true; and the explanation was presented as a separate sentence describing a potential explanatory account of those facts. For example, one set of facts read: “This object has a high-intensity light bulb and a pouch that can contain liquid”; and the associated explanation read: “This object was designed to kill bugs.”

For the categorization task, the facts and explanation were combined into a single sentence, following Ahn, et al. (2000)(e.g., “This object has a high-intensity light bulb and a pouch that can contain liquid because it was

designed to kill bugs.”). Of the 24 standards, 16 were test items, and 8 were fillers. The fillers were constructed to minimize the likelihood of demand characteristics in the categorization task (and will be described in greater detail in the context of that task). Of the 16 test items, 6 standards were taken without modification from the materials used in the Ahn, et al. (2000) study described above.

In addition to the six standards from Ahn et al.’s study, we constructed an additional ten items, five of which were designed to express good explanations; and five to express poor explanations. The good-explanation items were designed so that the explanation specified a general and plausible causal relation that accounted for all of the stated facts (as verified in the explanation-rating task described below). (See Table 2 for sample items). These additional ten standards were adapted from materials used in studies of category-based inference by Sloman (1994) and by Patalano, Chin-Parker, and Ross (2006).

The experiment consisted of two paper-and-pencil tasks, counterbalanced for order across participants: evaluating explanatory quality and carrying out a two-alternative forced-choice categorization task. For the explanation-rating task, participants were given all 24 standards, each arranged on a separate page of the test booklet. The facts were presented at the top of each page in a separate sentence, with the candidate explanation presented immediately below the facts. Participants were instructed to read the facts and the explanation, and then to provide a rating of how good or satisfying they found the explanation to be on a scale from 1 (very poor) to 7 (very good). They were then instructed to briefly explain their response.

The categorization task was identical to that of Experiment 4 in Ahn, et al. (2000) as described earlier. Specifically, it was a forced-choice task between two alternatives, in which participants were instructed to choose the one alternative that should be categorized with the target. Each triad—a standard together with both alternatives—was presented on a separate page of the test booklet, with the standard at the top of the page, and the alternatives below it. For the test items, one alternative shared the cause with the standard, but not the effects; the other alternative shared the effects, but not the cause. Sample items are presented in Table 2.

The eight filler items were constructed to vary the pattern so as to minimize the possibility of task demands. To this end, for four fillers, the alternatives both shared the cause with the standard, but differed on the effects; and for the remaining four fillers, both alternatives shared the effects, but differed on the cause.

For each participant, item order was randomized in both tasks; for the forced choice task, left-right presentation of alternatives was also randomized. Participants completed the tasks at their own pace: the categorization task required roughly 15 minutes to complete; the explanation-rating task, 25 minutes.

Results We submitted the proportion of responses for matching on the basis of a shared cause to a 2×2 analysis of variance (ANOVA) for a mixed-factorial design. There was no reliable interaction ($p = .28$), nor any reliable main effect of task order ($p = .25$): subsequent analyses are collapsed over this variable.

Table 2: Sample items from the study.

Standard	Shared Cause	Shared Effect
This person makes a good salesperson and persuasive public speaker, because she used to be a lawyer. (GOOD)	This person has a strong vocabulary and several enemies, because she used to be a lawyer.	This person makes a good salesperson and persuasive public speaker, because she likes to please people.
This person is estranged from his children and has few close friends, because he is a traveling salesman. (GOOD)	This person is outgoing and has many business contacts, because he is a traveling salesman.	This person is estranged from his children and has few friends, because he has a drinking problem.
This man has a six-figure income and works out regularly, because he was home-schooled as a child. (POOR)	This man has been divorced twice and is an avid moviegoer, because he was home-schooled as a child.	This man has a six-figure income and works out regularly, because he was the eldest of three siblings.
This person is a baseball fan and likes to read political magazines, because he is a night watchman. (POOR)	This person likes to play racquetball and watch documentaries, because he is a night watchman.	This person is a baseball fan and likes to read political magazines, because he is an actor.

In support of our main claim—that the size of the causal status effect depends on the goodness of the explanation in which the causal information is embedded—we found a reliable main effect of explanatory goodness, $F(1, 18) = 19.68, p < .001$. Planned comparisons reveal that the good explanations ($M = .77, SD = .19$) indeed exhibited a stronger causal status effect than the poor explanations ($M = .52, SD = .24$), $t(19) = 4.41, p < .001, d = 1.13$. That is, participants preferred to match on the basis of a shared cause to a much greater extent if that information was part of a good explanation.

The explanatory goodness ratings were in accord with the intended design: When participants rated the goodness of the explanations, the a priori “good” explanations were given higher goodness ratings ($M = 4.46, SD = .74$) than those designed to be “poor” explanations ($M = 2.86, SD = .75$), $F(1, 18) = 425.86, p < .001, d = 2.31$. (We have since

confirmed this finding with a separate group of participants who only rated the quality of the explanations).

Discussion The results of this study provide support for the main claim of this paper: that the causal status effect derives from the role that causal information plays in an explanation; and specifically, that the magnitude of the causal status effect depends on the goodness of the supporting explanation. This finding adds support to the position that explanatory structure is important in categorization (e.g., Keil, 1989; Lombrozo, 2006, 2007a; Murphy & Medin, 1985; Rips, 1989). According to this explanation-based perspective, categorization is treated as an inference to the best explanation. This perspective suggests that a primary research focus should be on investigating properties that characterize the quality of explanations, and the role that those properties play in cognitive tasks, such as categorization.

A possible objection to our study would be that we simply constructed explanations that were inconsistent with participants' prior knowledge. Ahn et al., (2000) found that the causal status effect can be eliminated if the stated causal relationship contradicts background knowledge. However, this consideration does not apply here, because we were careful to ensure that the poor explanations in our study did *not* contradict known facts. Rather, the poor explanations were designed to lack explanatory goodness. In some cases, there was no general causal connection between the facts and the explanation; in others, the explanation failed to explain all of the facts; and in others, there was more than one available explanation for the facts.

Conclusions The explanatory approach is compatible with, but broader than, the recent highly influential approaches that have focused strongly on causal knowledge—such as the causal status hypothesis (e.g., Ahn, 1998) and causal model theory (e.g., Rehder & Hastie, 2001). Causality-based theories tend to focus chiefly on characterizing causal knowledge and the causal reasoning processes that operate over it to support classification decisions. In contrast, the explanation-based approach focuses on characterizing the structure and quality of explanations and how they influence category processing. The explanatory approach often involves various kinds of causal structures, but it can also draw on other kinds of higher-order explanatory structures that do not rely on causality (e.g., information about perceptual or mathematical concepts, such as *symmetry* or *integral domain*)(Wilson & Keil, 1998). However, because causal information is critical for many kinds of mechanistic and functional explanations, the two approaches may turn out to be fellow travelers for most purposes. (See Lombrozo & Carey, 2006, for a recent proposal along these lines—the Explanation for Export hypothesis—which assigns a critical role to the nature of the causal relationship binding a cause to an effect).

At this point, it is fair to ask whether our efforts to explicate explanatory goodness are overkill. Arguably, a

direct focus on causal knowledge would be more parsimonious and would render the problem more tractable. For several reasons, we believe that a more direct focus on explanatory knowledge is needed. First, and most importantly, as Lombrozo (2006) notes, there are factors related to explanatory goodness that cannot be reduced to properties of causal knowledge, but which may be psychologically relevant for categorization. Second, although it might seem that a direct focus on causal knowledge would render the problem of centrality more tractable, recent theoretical and empirical work suggests that a consideration of explanatory goodness and explanation-based processes helps to focus research on the right questions (e.g., Lombrozo, 2006, 2007a; Lombrozo & Carey, 2006; Thagard, 1989; Pennington & Hastie, 1988). For example, using their influential explanation-based model of juror decision-making, Pennington and Hastie (1988) have shown that jurors arrive at a decision to acquit or convict a defendant on the basis of the most coherent explanation constructed for the evidence presented at trial. The success of their model demonstrates how a complex reasoning process can be profitably understood in explanation-based terms.

A third point is that a focus on categorization as inference to the best explanation may bring together areas of research that have been typically treated as separate. For example, it may be useful to think of analogy as another kind of inference to the best explanation. An analogy typically involves a mapping in which inferences are imported from a well-understood domain (the base) to one that is more poorly understood (the target)(Gentner, 1983). People intuitively prefer mappings that yield a good explanation, as evidenced by a tacit preference for systematicity—that is, a preference for depth and higher-order constraining relations in the resulting relational structure (e.g., Clement & Gentner, 1991). Though there have been arguments for a connection between analogy and categorization (e.g., Kuehne, et al., 2000; Ramscar & Pain, 1996), the relationship between them has not received much attention. Focusing more explicitly on the nature of explanatory processes may help to draw out the important commonalities, as well as meaningful differences, between categorization and other explanatory processes.

In summary, the results of our study offer evidence that the causal status effect depends on the quality of the explanation in which the causal information is embedded. These findings invite further exploration of the properties of explanation that influence categorization. Do simpler explanations (i.e., those that appeal to fewer causes; Lombrozo & Carey, 2006) lead to a stronger causal status effect than explanations that are more broad (i.e., account for more facts), or deeper? How important is depth of explanation in moderating the causal status effect? Do the same properties of explanatory goodness operate across a variety of cognitive processes (such as categorization and analogy), or to the same extent? Finally, focusing on explanatory structure may help to shed light on long-

standing debates over the specific role that background knowledge plays in categorization.

Acknowledgments

This research was supported by ONR Grant N00014-02-1-0078. We thank the Similarity and Cognition group for their valuable input.

References

- Ahn, W. (1998). Why are different features central for natural kinds and artifacts? *Cognition*, *69*, 135-178.
- Ahn, W., & Kim, N. S. (2000). The causal status effect in categorization: An overview. In D. L. Medin (Ed.), *The psychology of learning and motivation* (pp. 23–65). San Diego, CA: Academic Press.
- 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.
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: Plenum.
- Clement, C. A., & Gentner, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive Science*, *15*, 89-132.
- Forbus, K. D., & Gentner, D. (1986). Causal reasoning about quantities. *Proceedings of the Eighth Annual Meeting of the Cognitive Science Society*, 196-207.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155-170.
- Keil, F. (1989). *Concepts, Kinds, and Cognitive Development*. Cambridge, MA: MIT Press.
- Kuehne, S. E., Forbus, K. D., Gentner, D., & Quinn, B. (2000). SEQL - Category learning as progressive abstraction using structure mapping. *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*.
- Lombrozo, T. (2007a). Simplicity and probability in causal explanation. *Cognitive Psychology*, *55*, 232-257.
- Lombrozo (2007b). The effects of mechanistic and functional explanations on categorization. *Proceedings of the 29th Annual Conference of the Cognitive Science Society*.
- Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences*, *10*, 464-470.
- Lombrozo, T. & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, *99*, 167-204.
- Marsh, J. K., & Ahn, W. (2006). The role of causal status versus inter-feature links in feature weighting. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*.
- Medin, D. L. & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, *20*, 158-190.
- Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-196). New York, NY: Cambridge University Press.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289-316.
- Patalano, A. L., Chin-Parker, S., & Ross, B. H. (2006). The importance of being coherent: Category coherence, cross-classification, and reasoning. *Journal of Memory and Language*, *54*, 407-424.
- Pennington, N. and Hastie, R. (1988). Explanation-based decision making: The effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 521-533.
- Ramscar, M., & Pain, H. (1996). Can a real distinction be made between cognitive theories of analogy and categorization? *Proceedings of the 18th Annual Conference of the Cognitive Science Society*.
- Read, S. J., & Marcus-Newhall, A. R. (1993). Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology*, *65*, 429-447.
- Rehder, B. & Burnett, R. (2005). Feature inference and the causal structure of categories. *Cognitive Psychology*, *50*, 264-314.
- 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.
- Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21-59). New York: Cambridge University Press.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573-605.
- Rottman, B. M., & Gentner, D. (2006). Expertise effects on sorting strategies of causal phenomena. *Proceedings of the 28th Annual Meeting of the Cognitive Science Society*.
- Salmon, W. (1989) *Four Decades of Scientific Explanation*. University of Minnesota Press
- Slovan, S. A. (1994). When explanations compete: The role of explanatory coherence on judgments of likelihood. *Cognition*, *52*, 1-21.
- Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, *12*, 435-467.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327-352.
- Wilson, R.A. and Keil, F.C. (1998). The shadows and shallows of explanation. *Minds and Machines*, *8*, 137-159.