Proceedings of the Interdisciplinary Workshop on Categorization and Similarity, University of Edinburgh, 115-121.

# Features of Similarity and Category-Based Induction

**Evan Heit**
Department of Psychology
University of Warwick
Coventry CV4 7AL
United Kingdom
E.Heit@warwick.ac.uk

**Abstract**

A classic feature-set model of similarity, the contrast model of Tversky (1977), is applied to a range of phenomena in category-based induction. These phenomena include basic similarity effects, typicality and asymmetry effects, diversity effects, effects of projectibility of properties, and contextual influences on similarity. These analyses help to extend the contrast model to a new area of research as well as to place constraints on the model itself.

Although categorization research has largely focused on people's ability to infer taxonomic category labels, categories facilitate a number of cognitive abilities and functions. One of our most important abilities is inductive inference (Anderson, 1991; Billman & Heit, 1988; Heit, 1992; Osherson, Smith, Wilkie, Lopez, & Shafir, 1990), and category-level information enables a rich set of inferences. For example, you might not be able to infer much about Peter until you are told that Peter is a goldfish, in which case you could infer that he will swim and he will eat fish food. Indeed, these inferences too can be thought of as categorization judgments, such as putting Peter into the *swimmers* category. What makes category-based induction especially powerful is our capacity to project information from one category to another, similar category. For example, consider the following inductive argument.

```
Tunas thrive in sunlight     (1)
--------------------------
Goldfish thrive in sunlight.
```

If the premise above the line is known to be true, then the conclusion below the line is plausibly true, in light of the similarity between the premise category *tunas* and the conclusion category *goldfish*. Underlying this assumption is Mill's (1874) thesis that "because *a* resembles *b* in one or more properties, that it does so in a certain other property." In other words, because tunas and goldfish are similar in some respects, it seems plausible that they will be similar in terms of a novel property, *thrives in sunlight*, as well.

Given this close link between similarity and induction, it is natural to ask whether psychological models of similarity can account for the key findings in induction. In this paper, Tversky's (1977) contrast model, one of the classic models of similarity, will be applied to a number of results in induction. The contrast model is one of the simpler feature-set accounts of similarity, which together represent the main alternative to geometrical or multidimensional models of similarity (e.g., Shepard, 1974). The contrast model has not been previously applied in the area of category-based induction, so this paper will have a somewhat exploratory character. How well can a very successful model of similarity account for induction data? What additional operating assumptions must be made? In other words, the results on induction will be used to extend the contrast model to another area as well as to place additional constraints on the model itself.

In the next section of this paper, Tversky's contrast model will be reviewed, and in following sections, several of the main phenomena in induction will be presented and discussed. For the most part, previous applications of similarity models--and categorization models--have not addressed these phenomena. However, in light of the close links between similarity and induction, as well as the centrality of these findings to induction, it is hoped that closer attention to results in induction will be useful in developing better models of similarity.

## The Contrast Model of Similarity

The contrast model relies on featural representation of objects, and it is used to compute the similarity between the representations of two objects. Similarity is defined as an increasing function of common features, that is features in common to the two objects, and as a decreasing function of distinctive features, that is features that apply to one object but not the other. This definition is formalized in the following equation.

$$S(a,b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$$

The similarity, S, between *a* and *b*, is defined in terms of the features common to *a* and *b*, $A \cap B$, the features that are distinctive to *a*, A - B, and the features that are distinctive to *b*, B - A. The variables $\theta$, $\alpha$, and $\beta$ are non-negative valued free parameters that determine the relative weights of these three components of similarity. These parameters will be discussed in further detail as the model is applied to results in induction, but here it is noted that these parameters allow the model some flexibility, for example, in whether common or distinctive features will have more influence. Additionally, the function *f* measures the salience of a particular set of features. According to Tversky, the salience of a feature depends on two factors, intensity and context. The intensity of a feature refers to its physical salience as well as other inherent, stable aspects. For example, a large red chimney on a house would probably be more salient than a small gray one. Also, Tversky recognized that the salience of features is not fixed but may vary with context or goals (see also Medin, Goldstone, & Gentner, 1993). For example, chimneys might be especially salient to someone who repairs chimneys.

Another useful quality of the contrast model is that it can allow for different roles in similarity comparisons, to account for possible asymmetries. For example, in the statement "Twiglets are similar to pretzels," *twiglets* is the subject and *pretzels* is the referent. In applying the contrast model to such statements, Tversky assigned the subject to the *a* term and the referent to the *b* term. Here it will be assumed that the premise category of an inductive argument is the *a* term and the conclusion category is the *b* term. Also, it is noted that the contrast model is quite general and could be applied to features at various levels of description, such as perceptual features or semantic features. Tversky suggested that the model would be applied to lists of features following processes of extraction and compilation (p. 329). That is, there is no assumption that objects are permanently represented as straightforward feature lists, but rather that people can compile such lists when needed.

To apply the contrast model to inductive inference, it will be assumed that there is a monotonic relation between the similarity between *a* and *b*, S(a,b), and the strength of an inductive argument of the form "Category *a* has Property P; therefore Category *b* has Property P." In experimental situations, there have been several measures of inductive strength. Perhaps the most direct method is asking subjects to rate the strength of various arguments, or choose which of two alternative arguments is stronger. Alternately subjects will estimate the probability of an argument's conclusion being true, which may be a more familiar kind of judgment compared to assessing inductive strength.

## Phenomena in Category-Based Induction

A number of papers have established a core set of findings in induction that any model would need to address. The seminal paper was by Rips (1975), but it only addressed rather simple inductive arguments, i.e., with one premise category and one conclusion category. Work by Osherson et al. (1990) documented a wider range of phenomena, including more complex arguments with multiple premise categories. These papers focused on the effects of categories on induction, such as effects of similarity between premise and conclusion categories, rather than the influences of the property itself that is projected (such as *thrives in sunlight* in Argument 1). However, other work (e.g., Gelman, 1988; Heit & Rubinstein, 1994) has documented some property effects, which could be especially challenging for a model that mainly addresses similarity between objects or categories. Next, a selection from this core set of results will be presented, to some extent moving from the simpler to the more complex phenomena.

### Similarity Effects

The most widespread and robust finding in induction is a basic similarity effect. For example, "Cows have Property P, therefore Sheep have Property P" is a stronger argument than "Cows have Property P, therefore Hedgehogs have Property P" (Osherson et al., 1990; Rips, 1975). People are more willing to project a property of cows to a similar category (*sheep*) than to a less similar category (*hedgehogs*).

This result would not seem to place too many constraints on a model of similarity; all that is needed is an assumption that judgments of argument strength, or judgments of the likelihood of the conclusion, will be monotonically related to the similarity between the premise category and the conclusion category (see also Rips, 1975). This result does not seem to lead to any other constraints on the model, such as constraints on the parameters θ, α, and β. That is, the basic finding that inductive strength is well-correlated with similarity does not indicate the relative contributions of common features, distinctive features of *a*, and distinctive features of *b*. Indeed, it would be interesting in future research to look at relative contributions of common and distinctive features, by manipulating these variables separately. Heit (in press) has recently argued that what is critical in inductive reasoning is the number of distinctive features of the premise category, that is, the idiosyncrasy of the premise category, but common features may well also make a contribution.

Finally, it is useful to mention a limiting-case variant of similarity effects, referred to by Osherson et al. (1990) as premise-conclusion identity. Namely, an argument such as

```
Tunas thrive in sunlight     (2)
-------------------------
Tunas thrive in sunlight,
```

with identical premise and conclusion categories, is at least as strong as any other argument. This result maps well onto the property of minimality for a model of similarity, that S(x,x) = S(x,y) for any x and y.

### Typicality and Asymmetry Effects

Demonstrations of typicality effects in induction indicate that when it comes to premises, not all categories are created equal. Some categories, when used in premises in an inductive argument, have a greater impact than other categories. To used Goodman's (1955) term, it could be said that typical categories have greater *entrenchment* than atypical categories (see also Rips, 1975; Shipley, 1993). Consider the following two arguments:

```
Cows have Property P         (3)
-------------------------
Hedgehogs have Property P
```

and

```
Hedgehogs have Property P    (4)
-------------------------
Cows have Property P.
```

The first argument, with a typical premise category, seems stronger than the second, with an atypical premise category. Facts about atypical categories just don't seem to generalize too well. There is considerable evidence that typicality effects are widespread. For example, Rips (1975) looked at a large number of inductive arguments involving pairs of mammals or birds, and found that the typicality of the premise category was a significant, independent predictor of the overall strength of the argument, and notably, the typicality of the conclusion category did not predict argument strength.

How can this result be used to constrain the contrast model? In general, the model can predict asymmetry in two ways, either with $\alpha > \beta$ or $\beta > \alpha$. That is, either the distinctive features of the *a* term or the distinctive features of the *b* term will play a greater role. Note that if $\alpha = \beta$, then S(a,b) must equal S(b,a) and hence there will be no asymmetry. Some additional results can be used to decide between the two possibilities, either $\alpha$ higher or $\beta$ higher. Osherson et al. (1990) presented a phenomenon called *premise-conclusion inclusion*, in which arguments are perfectly strong when the premise category includes the conclusion category. Arguments 5 and 6 illustrate this phenomenon.

```
All Mammals have Property P  (5)
---------------------------
Cows have Property P
```

and

```
All Mammals have Property P  (6)
---------------------------
Hedgehogs have Property P.
```

According to Osherson et al. (1990), these two arguments are perfectly strong, and hence equally strong. This equality places a strong constraint on the contrast model, namely, S(All Mammals, Cows) = S(All Mammals, Hedgehogs) = S (All Mammals, *m*), where *m* is any mammal. Cows, hedgehogs, and other mammals would all have their own, different distinctive features, so for this equality to be true, the $\beta$ parameter must be set at 0. Hence, distinctive features of the conclusion category must have no impact at all on induction. This constraint is rather severe, and, again, it would be useful to test it experimentally by explicitly varying the distinctive features of the conclusions of inductive arguments. However, this predicted lack of sensitivity to distinctive features of the conclusion category does seem to fit with the Rips (1975) finding that typicality of the conclusion category is not correlated with inductive strength.

Now if $\beta = 0$, then for there to be any asymmetry, it must be the case that $\alpha > 0$. What other constraints can be derived? Returning to Arguments 3 and 4, the asymmetry, S(Cows, Hedgehogs) > S(Hedgehogs, Cows), can only hold when the distinctive features of hedgehogs are more salient than the distinctive features of cows (Tversky, 1977). One way to think of this assumption is that cows are represented as typical mammals, with few distinctive features, but hedgehogs are represented as being highly distinctive from other mammals, including cows. However, this assumption would be a bit tricky to evaluate empirically, because a simple feature listing task might not reveal the salience of various features for induction. Indeed, Heit (in press) suggested that people's inductive inferences could be influenced by "placeholder features," that is, features correponding to a deeply-held but non-specific belief. For example, people might have a strong belief that there is something highly distinctive about hedgehogs, but it might not be easy to list this something as a feature on a feature listing task (cf., Medin & Ortony's, 1989, account of psychological essentialism).

A further twist is that in some analyses of asymmetry, Tversky (1977) assumed that prototypical items would have more salient features than other items, e.g., the number 100 is more prototypical than the number 103, and it is more salient as well. This assumption may seem to conflict with the idea above that atypical items will have especially salient distinctive features. One complication is distinguishing distinctive features from familiarity: 100 is possibly more familiar than 103, but it is not obvious that 100 has more distinctive features. Also, items that were prototypical in Tversky's analyses (e.g., the number 100, a pure example of the color red) would not necessarily make strong premises in inductive arguments.

In conclusion, the contrast model can be applied to typicality and asymmetry effects in induction, and these results yield useful constraints on the model. However, the success of the model does depend on how objects are represented, i.e., that atypical items must have more salient distinctive features than typical items. The evidence supporting this assumption is not yet clear.

## Diversity Effects

One of the fascinating findings in inductive reasoning is that while a typical category can make a strong premise, two typical categories together can make a worse case than one typical category and one atypical category together. Consider the following arguments:

```
Cows have Property P         (7)
Horses have Property P
---------------------------
All Mammals have Property P
```

and

```
Cows have Property P         (8)
Hedgehogs have Property P
----------------------------
All Mammals have Property P.
```

People tend to find arguments such as Argument 8 stronger; arguments with premises that contain a diverse set of categories are favored over arguments that contain a narrow range of categories (Osherson et al., 1990). In the present example, cows and horses together make a weaker case even though horses are more typical than hedgehogs. Intuitively, cows and horses are so similar that finding out that horses have some property adds little information when you already know that cows have the property. That is, Argument 7 would hardly be any stronger than Argument 9, below. On the other hand, getting information about hedgehogs seems to tell you something that you don't already know, so Argument 8 would be quite a bit stronger than Argument 9.

```
Cows have Property P        (9)
---------------------------
All Mammals have Property P
```

The first question considered is how to apply the contrast model to arguments with multiple premises, such as Arguments 7 and 8. The S function takes only two variables. Somehow, knowledge about the multiple categories in the premise must be combined. One solution is to treat the set A as the *intersection* of features of the premise categories. To evaluate, say, Argument 7, a person would form a representation of what cows and horses have in common. Then the S function could be used to assess the feature overlap between this representation and a representation of the conclusion category, mammals. The intersection of features for cows and horses would seem to reflect information about large, domestic mammals, often found on farms living in herds--not much different than the features of cows taken alone. Hence the evaluation of Argument 7 would not be much different than the evaluation of Argument 9.

Now consider Argument 8. What is the intersection of features for cows and hedgehogs? Cows and hedgehogs are both mammals, of course, so anything that is true of all mammals would be true of both cows and hedgehogs. For example, cows and hedgehogs both have skin, warm blood, and teeth. But cows and hedgehogs do not have much else in common. It is proposed that the intersection of diverse categories will tend to have few idiosyncratic or distinctive features. The features that diverse categories have in common will tend to be features that are shared by many other categories as well. Hence, S(Cows $\leftrightarrow$ Hedgehogs, All Mammals) should be quite high, because the intersection of cows and hedgehogs will have many common features with all mammals, and few distinctive features. Thus the inductive strength of Argument 8 would be considerably greater than that of Argument 7 or Argument 9.

Thus, this simple extension of the contrast model, in which the intersection of features of all the premise categories is assigned to the first position of the S function, seems promising as a way to account for the diversity effect. It does seem plausible that the intersection of features of similar categories would tend to resemble the feature sets of each category taken alone, whereas the intersection of features of diverse categories would tend to have general features that are common to other categories as well. However, it would be useful to evaluate this assumption empirically, perhaps using feature-listing tasks.

A final issue of interest is that while diversity effects do seem to be robust in many situations (see also, Lopez, 1993; Osherson, Stern, Wilkie, Stob, & Smith, 1991), it seems that young children quite consistently do not show diversity effects (Gutheil & Gelman, 1997; Lopez, Gelman, Gutheil, & Smith, 1992). It would be interesting to develop, within the framework of the contrast model, an explanation for why children do not show diversity effects. Perhaps the operation of combining multiple premise categories and finding the intersection of their features is too difficult for children in an experimental setting.

## Non-Monotonic Effects

Although it would seem from the preceding section that diverse premise categories, such as *cows* and *hedgehogs*, promote induction, it is also the case that too much diversity can hurt induction. Consider the following example, from Osherson et al. (1990).

```
Flies have Property P        (10)
---------------------
Bees have Property P

Flies have Property P        (11)
Orangutans have Property P
--------------------------
Bees have Property P
```

Interestingly, people find Argument 10 to be stronger than Argument 11, even though Argument 11 includes a second premise that adds diversity. The critical difference in this example is that orangutans come from a different superordinate category (*mammals*) than flies and bees (*insects*), whereas in the previous section, cows and hedgehogs are contained in the same superordinate category, *mammals*.

In terms of the contrast model, Argument 11 is particularly weak because the intersection of fly features and orangutan features is very small, much smaller than the set of known features of flies. When the S function is used to assess the overlap between the premises, flies and orangutans, and the conclusion, bees, there will be very few common features remaining. Hence the similarity between the representation of the premises and the representation of the conclusion will be very low, and likewise inductive strength will be low. In contrast, for Argument 8, cows and hedgehogs do have some features in common, and indeed these features are those that are true of all mammals. Thus the conclusion that all mammals have the property is well-supported by these diverse premises.

The main point to be drawn from this section is that the contrast model is able to account for non-monotonic effects as well as diversity effects, according to the rough idea that too much diversity will reduce common features so much that inductions are weak. Whether the model predicts non-monotonicity or diversity for a particular set of categories would depend on the representations of the categories.

## Projectible Versus Non-Projectible Properties

The preceding examples have focused on how induction is influenced by the categories in an inductive argument, e.g., whether there is a typical category in the premise such as cow or an atypical category such as hedgehog. However, just as all categories are not created equal in inductive reasoning, not all properties are equal either: People reason differently depending on what Property P actually is. Some previous work (e.g., Rips, 1975; Osherson et al., 1990), while documenting a number of interesting category-based phenomena, has only looked at fairly unfamiliar, or "blank," anatomical properties of animals, such as "has sesamoid bones" or "has BCC in its blood." For such unfamiliar properties, it seems plausible that the property itself has little influence on reasoning. But inductive reasoning with blank properties is only one facet of inductive reasoning more broadly considered, in which the property itself plays an important role. The remaining examples in this paper show some influences of properties on induction. These examples provide important challenges for the contrast model and indeed for any model applied to induction.

The first property-based phenomenon to be described is a classic in inductive reasoning. Goodman (1955), in posing his riddle of induction, noted that some properties are more *projectible*, or more easily projected, than other properties. Some properties seem particularly transient or idiosyncratic, and they seem unlikely to project from one category or individual to another (see also Nisbett, Jepson, Krantz, & Kunda, 1983). To use an example from Gelman (1988), imagine that you see a rabbit eating alfalfa. You can probably conclude that other rabbits would like to eat alfalfa as well. But if you see a rabbit that is dirty, you probably would not conclude that other rabbits are dirty as well. In other words, Argument 12 is stronger than Argument 13.

```
Rabbit 1 eats alfalfa        (12)
---------------------
Rabbit 2 eats alfalfa

Rabbit 1 is dirty            (13)
-----------------
Rabbit 2 is dirty
```

A simplistic application of the contrast model would not predict a difference in the strengths of these two arguments, as long as the similarity between Rabbit 1 and Rabbit 2 stays the same. But the contrast model is flexible in terms of its $f$ function, which determines the salience of various features. To apply similarity-based reasoning to inductive inferences about a range of properties, a person would need to know which features should be salient or relevant for various inferences (Heit & Rubinstein, 1994). Goodman (1955) suggested that we have abstract beliefs, called *overhypotheses*, describing the scope of properties such as food preferences and transient properties (see also Shipley, 1993). An example of an overhypothesis would be "different kinds of animals eat characteristic foods." Such overhypotheses could be used to constrain the salience function $f$. For example, to make an inference about food preferences, features related to species of animal would be highly salient. In this context, Rabbit 1 and Rabbit 2 would be highly similar in Argument 12, and this argument would be very strong. For a property such as *is dirty* in Argument 13, the overhypothesis might be that such an idiosyncratic and transient property would depend on very localized circumstances, such as whether an animal is sitting in a puddle or pile of mud. Species-level features would not be particularly salient in function $f$, hence Rabbit 1 and Rabbit 2 would not seem very similar and the argument would not seem very strong.

This discussion serves to illustrate that the contrast model is flexible enough to allow for differences due to projectible versus non-projectible properties. It would be incorrect to say that the contrast model predicts no sensitivity to kind of property. Indeed, Tversky (1977) discussed a number of factors that might affect the salience function, such as diagnosticity of features in the context of other objects that are present (see also Medin et al., 1993). However, at present there is no complete account of how properties in inductive arguments would affect the function $f$. It does seem plausible that people have overhypotheses or other knowledge about feature dependencies, so it does seem reasonable to suggest that people indeed would be able to vary the $f$ function according to the context of an inductive argument.

## Different Kinds of Similarity

Although it might seem from the previous section that some properties (the projectible ones) are good for induction and other properties (the non-projectible ones) do not promote induction, the picture is actually more complicated and more interesting. Depending on the argument, that is, depending on the categories in an inductive argument, a particular property may be projectible or non-projectible or somewhere in between.

Consider the following example, from Heit and Rubinstein (1994). For a typical blank anatomical property, such as "has a liver with two chambers," people will make stronger inferences from chickens to hawks than from tigers to hawks. Because chickens and hawks are from the same biological category, and share many internal properties, people are quite willing to project a novel anatomical property from one bird to another. But since tigers and hawks differ in terms of many known internal biological properties, it seems less likely that a novel anatomical property will project from one to the other. This result illustrates the priority of biological categories that has been observed in induction (e.g., Carey, 1985; Gelman, 1988).

However, now consider the behavioral property "prefers to feed at night." Heit and Rubinstein (1994) found that inferences for behavioral properties concerning feeding and predation were weaker between the categories *chicken* and *hawk* than between the categories *tiger* and *hawk*--the opposite of the result for anatomical properties. Here, it seems that despite the considerable biological differences between tigers and hawks, people were influenced by the known similarities between these two animals in terms of predatory behavior, thus making strong inferences about a novel behavioral property. In comparison, chickens and hawks differ in terms of predatory behavior (with chickens tending to be pacifists), thus people were less willing to project a novel behavioral property between these two animals. Putting together these results, it seems that each property is more projectible for a different pair of animals. It is not simply the case that some properties are always more projectible than other properties.

This pattern of results can be addressed in a straightforward manner using the contrast model. As in the examples in the previous section, a person would have to apply knowledge about feature dependencies to constrain the salience function *f*. For ordinary anatomical properties such as "has a liver with two chambers," the salience function could emphasize other anatomical features referring to insides of animals. Hence, the model would predict that inferences would be stronger from hawks to chickens (both birds, of course) than from tigers to chickens.

In contrast, when reasoning about a novel behavioral property such as "prefers to feed at night," a person would give additional salience to other behavioral features. The assumption is that animal behaviors will be particularly interpredictive. For example, a person might consider known facts concerning whether various species prey on weaker animals, and whether they use sharp claws for attacking, to derive an estimate of similarity. On this basis, hawks and tigers would seem particularly similar, and hawks and chickens would seem less similar, hence an inductive argument with a behavioral property would be stronger for hawks and tigers than for hawks and chickens.

Again, this account of the use of different kinds of similarity information serves to illustrate the flexibility of the contrast model. Although this model does not actually include a mechanism for varying the salience function *f* based on the property under consideration, it does seem plausible that people would be able to do so. Also, such a flexible conception of similarity would be essential to any successful account of induction: A model of induction that depends on a fixed notion of feature salience would not be able to account for the Heit and Rubinstein (1994) findings.

## Discussion

### Summary

Although the contrast model of similarity was not designed with results on induction in mind, it does seem that the model has some promise in this area. Possibly most interesting is what has been learned about the contrast model in applying it to these results. The analysis of typicality and asymmetry effects suggests that distinctive features of the conclusion category should have no effect on induction, whereas the distinctive features of the premise category should have a significant effect. However, the relative impact of common features, compared to distinctive features of the premise, has not been clarified. The analysis of diversity effects suggests a way of assessing the similarity between more than two objects, such as in Argument 7, the similarity between cows and horses on the one hand and mammals on the other hand. Namely, it is possible to use intersections of category representations. The analyses of projectibility effects and of use of different kinds of similarity led to further constraints on the contrast model, in particular on the salience function *f*. Although Tversky (1977) did make some suggestions in his paper, additional constraints are needed for application to inductive arguments. More generally, these analyses have pointed to the importance of people's background knowledge in assessing similarity and performing category-based induction (see also Heit, 1997).

### Other Models of Induction

One way of evaluating the success of the contrast model here is to compare it to other models that were specifically designed to apply to induction. The first model, by Rips (1975), was just a regression model that predicted the strength of single-premise arguments from similarity as well as the typicality of the premise category. This model was applied only to similarity and typicality effects. A more elaborate model was presented by Osherson et al. (1990), and this model was applied to a wide range of phenomena, including similarity, typicality, and diversity effects. The Osherson et al. model also has feature-set representation, but it uses a more complex similarity function and it makes further computational assumptions. It appears that the contrast model can account for all the phenomena listed by Osherson et al., but further analysis will be required.

A connectionist model by Sloman (1993) addresses many of the same phenomena as the Osherson et al. (1990) model. Interestingly, the Sloman (1993) model makes just the opposite assumption of the contrast model, namely that distinctive features of the conclusion category should dominate rather than distinctive features of the premise categories. The Sloman (1993) model would seem to have trouble explaining the classic Rips (1975) result that premise typicality is correlated with inductive strength but conclusion typicality is uncorrelated with inductive strength. However, it should also be pointed out that the contrast model would have difficulty explaining Sloman's apparent contradictions to the premise-conclusion inclusion phenomenon (see Arguments 5 and 6) above. Clearly, further empirical work is needed to assess the roles of distinctive features of premises and conclusions, and the assumption made for the contrast model that distinctive features of the conclusion category have *no* influence may need to be weakened slightly.

Compared to the Osherson et al. (1990) and Sloman (1993) models, the contrast model is much simpler, in terms of mathematical operations required. Although parsimony of expression is not the only criterion for discriminating among formal models, the relative simplicity of the contrast model does raise the issue of when the additional complexities of the other models are truly needed. Because the contrast model is an established model of similarity, perhaps it should be considered a baseline model, and another, more complex model should be accepted only when distinctive evidence is provided for such a model over the contrast model. Also, the other models were not applied to arguments where the property made a difference. That is, the other models were not applied to projectibility effects or uses of different kinds of similarity as in Heit and Rubinstein (1994). However, these models could possibly adopt the present approach, namely allowing the salience of features used in similarity computations to vary depend on the property being inferred.

Finally, there are a number of parallels between this application of the contrast model and a Bayesian model of induction proposed by Heit (in press). In brief, the Bayesian model puts together background knowledge about distributions of properties with the information contained in the premises of the argument, to derive an estimate of the likelihood of the conclusion being correct. The Bayesian model accounts for the same phenomena described in this paper in terms of the contrast model, with a number of points of correspondence. For example, when an argument with multiple premises is evaluated, the Bayesian model uses background knowledge of properties that are true of all the premise categories. Likewise, in the contrast model, the similarity assessment is based on the intersection of properties of all the premise categories. The parallels between the Bayesian model and the contrast model are quite encouraging, because they suggest that many of the assumptions and predictions of the contrast model can be derived independently from normative Bayesian statistical theory.

## Conclusion

Although this initial application of the simple and elegant contrast model to induction has been encouraging, further work is needed in two areas. First, experimental work would be helpful to test the various assumptions and predictions of the model. Second, a much longer-term goal is to develop a better account of the function $f$, that is, developing a better account of how people decide which features are relevant to predicting other features (for further discussion, see Heit, in press). A better specification of the function $f$ might be considered the Holy Grail of induction, because it is needed to allow the contrast model as well as other accounts of induction to explain how people apply their knowledge about categories and properties to perform a wide range of inductive inferences.

## Acknowledgments

I am grateful to Ulrike Hahn, Douglas Medin, Steven Sloman, and David Over for comments on this research.

## References

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

Billman, D., & Heit, E. (1988). Observational learning without feedback: A simulation of an adaptive method. *Cognitive Science, 12,* 587-625.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: Bradford Books.

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

Goodman, N. (1955). *Fact, fiction, and forecast*. Cambridge, MA: Harvard University Press.

Goodman, N. (1972). *Problems and projects*. Indianapolis: Bobbs-Merrill.

Gutheil, G., & Gelman, S. A. (1997). Children's use of sample size and diversity information within basic-level categories. *Journal of Experimental Child Psychology, 64*, 159-174.

Heit, E. (1992). Categorization using chains of examples. *Cognitive Psychology, 24,* 341-380.

Heit, E. (1997). Knowledge and concept learning. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts, and categories*. London: Psychology Press.

Heit, E. (in press). A Bayesian analysis of some forms of inductive reasoning. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition*. Oxford University Press.

Heit, E., & Rubinstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 411-422.

Lopez, A. (1993). The diversity principle in the testing of arguments. *Memory & Cognition, 23*, 374-382.

Lopez, A., Gelman, S. A., Gutheil, G., & Smith, E. E. (1992). The development of category-based induction. *Child Development, 63,* 1070-1090.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254-278.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. Cambridge: Cambridge University Press.

Mill, J. S. (1874). *A system of logic*. New York: Harper.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review*, 339-363.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97,* 185-200.

Osherson, D. N., Stern, J., Wilkie, O., Stob, M., & Smith, E. E. (1991). Default probability. *Cognitive Science, 15*, 251-269.

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

Shepard, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika, 39,* 373-421.

Shipley, E. F. (1993). Categories, hierarchies, and induction. In D. L. Medin (Ed.), *The Psychology of Learning and Motivation* (Vol. 30). Academic Press.

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

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.