



ELSEVIER

Cognition 65 (1998) 87–101

COGNITION

Similarity as an explanatory construct

Steven A. Sloman^{a,*}, Lance J. Rips^b

^a*Department of Cognitive and Linguistic Sciences, Brown University, Box 1978,
Providence, RI 02912, USA*

^b*Psychology Department, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208, USA*

Abstract

Theories can be found throughout cognitive science that give an explanatory role to similarity. Such theories can be contrasted with those that model thought using abstract rules. We lay out four possible explanatory roles for similarity. We then review the computational pros and cons of similarity- and rule-based models and outline the empirical work that speaks to the psychological plausibility of the two frameworks. We conclude that an adequate model of human thought must take advantage of both the flexibility of similarity-based inference and the compositionality and certainty associated with rule-based inference. © 1998 Elsevier Science B.V.

Keywords: Similarity; Rule-based inference

1. Introduction

Poor old Aunt Bess has passed away, leaving you some items from her antique collection. There are the fondu forks and the candy dispensers. There's the well-thumbed copy of Hilgard and Marquis, and there's a dusty painting of a vase in multiple perspective. One of Uncle Herbert's better efforts? Or did Aunt Bess have a piece of art on her hands whose value she never realized?

Braque or Uncle Herb? This is a question about categorization: the painting's properties are given to you through perception, and your job is to assign it to a category. But you could also face a somewhat similar question about the inferences you are willing to make on the basis of your knowledge of the new painting. For example, suppose you see that this very painting has rigidium-based paint. The

* Corresponding author. e-mail: Steven_Sloman@brown.edu

question this time is: Do Picassos have rigidium-based paint? Do Whistlers have it? Finally, we can phrase some of the same issues in terms of decision making. Let's say this time that you're shopping for a painting with certain ideal properties, say, having some historical significance, being relatively inexpensive, and fitting the color scheme of your breakfast nook. Would you purchase this painting?

There's a simple theory of how people make all these judgments, a theory that has an enormous appeal to many cognitive scientists. According to this theory, the judgments crucially involve determining the similarity between the specimen and some other relevant entity and then basing the judgment on the resulting degree of similarity. In the case of categorizing, you determine the similarity between the specimen and the prototypes or examples of the categories to which it might belong. For example, you measure the similarity between the new painting and Braque's and between the new painting and Uncle Herbert's. If the similarity to Braque's is sufficiently great, you will call it a Braque; if the similarity to Uncle Herbert's is sufficiently great, you'll call it an Uncle Herbert; and if the similarity falls below the criterion for both categories, you'll place the painting in a new category of its own. In the case of inference-making, similarity may also be the operative principle. If the new painting has rigidium-based paint and if the painting is sufficiently similar to Picassos, then you may be tempted to infer that Picassos have rigidium paint too. And in the decision-making example, if the painting is sufficiently similar to your ideal piece of art, then you'll decide to purchase it.

Part of the appeal of this picture is that it fits so well the traditional world view of experimental psychologists. Explanation in psychology has at its disposal a few key concepts that it wields to account for a host of phenomena. Frequency is one such concept, salience is another, similarity is a third. If we can reduce an effect to a difference in frequency, salience, or similarity, then no more need be said. These three abide; but the greatest of these is similarity.

Or is it? Although similarity is appealing because of its wide range of application, it strikes many contemporary researchers as a little too easy. For one thing, your judgment of the similarity between two objects can fluctuate with changes in the task you're performing. Your assessment of the similarity between A and B may depend on the set of other objects in the stimulus set, on the direction of the comparison (*How similar is A to B?* versus *How similar is B to A?*), on whether the comparison is based on relational or nonrelational information, and probably on many other factors (Tversky, 1977; Goldstone et al., 1991). For another, explanations based on similarity often look circular. With a little brain storming, we can almost always come up with properties that A shares with B, for any B. With respect to these properties, A and B will be similar, but with respect to some other set of properties dissimilar (Goodman, 1955; Murphy and Medin, 1985). Suppose we want the new painting in our example to be a Braque. If we use similarity to decide the issue, we can rig this outcome by selecting or emphasizing properties that it shares with Braques, but we could have equally well rigged the decision in favor of Uncle Herb.

Because of these problems with similarity, we seem to have a choice. We can preserve similarity as an explanatory principle by tinkering with it to avoid its undesirable properties. Or we can give up on similarity, trying to handle categoriza-

tion, inference, decision making, and other cognitive skills through different mechanisms. Similarity works best in domains where there is little choice about the properties that determine the outcome. It seems most suspect in domains where the key properties are variable and task dependent. To get similarity to work in these higher-level domains, you need an independent way to select the relevant properties for input to the similarity process, but it often seems that these selection routines are doing all the important work. Why not abandon similarity altogether, then, and elaborate the selection routines?

Here's an example of the difficulty we have in mind (from a thought experiment that Tienson, 1988, has expounded): Take a triangle and remove a tiny piece from one vertex, replacing the piece with a straight line segment, as shown in Fig. 1. The result of this operation, on the right of Fig. 1, looks an awful lot like—we might even say, is awfully similar to—a triangle, but it isn't. It's a quadrilateral. To explain why we classify the figure as a quadrilateral rather than a triangle, we could say that we've computed the similarity of the object to quadrilaterals using the property of having four sides. But the similarity comparison is doing no useful work here. Once we know it has four sides, we have all we need to classify the figure correctly. Moreover, our untutored similarity judgments seem to reach a verdict opposite that of the classification decision. We might very well judge the figure more similar to triangles than to quadrilaterals.

Similarity as an explanation isn't very tempting in cases like that of Fig. 1, where we have theoretical information on which to base our decision. The same is true of the painting example, since whether the thing is a Braque or an Uncle Herbert surely depends on who painted it and not on its similarity to other paintings by these masters. Investigators who are suspicious of the role similarity plays in cognitive science have tried to capitalize on these cases, showing that theory also enters into more everyday instances of classification and reasoning. There are many different cognitive mechanisms we can choose from in order to carry out these theory-based decisions, but for convenience we will lump them under the heading of *rules*.

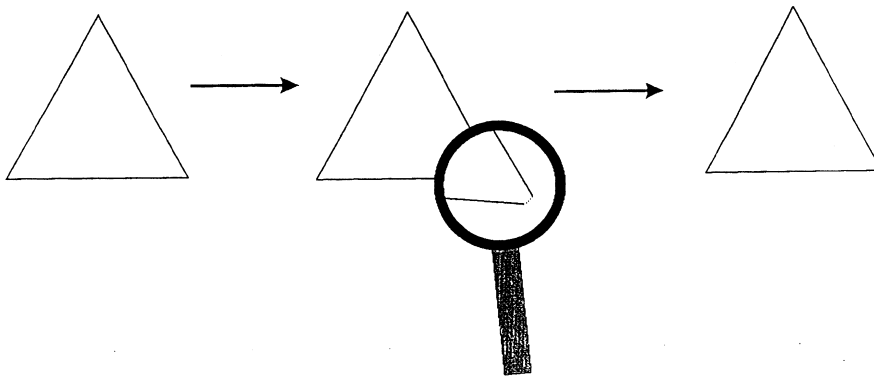


Fig. 1. Steps in Tienson's (1988) thought experiment: a small part of the vertex of the triangle at left is snipped out and replaced by a straight line. The resulting figure at right is a quadrilateral, despite retaining similarity to triangles.

The issue of general interest to cognitive science is whether rules or similarity or both are psychologically plausible ways to explain cognitive processes. Both kinds of systems have specific computational advantages and disadvantages and both kinds of systems find empirical motivation. This special issue considers the issue from several perspectives. Our papers span cognitive and developmental psychology, neuroscience, and theory of computation, and they include new empirical results as well as new efforts at theoretical integration. The papers certainly do not speak with a single voice; rather, they cover the range of problems and issues that animate the rules vs. similarity debate.

2. Four views of similarity

The role of similarity is a matter of debate in most areas of cognitive science that study thinking. Similarity has entered into accounts of concept formation, reasoning, decision making, and many other areas. The term typically refers to the outcome of a comparison among entities, usually a comparison based on many of the entities' properties. Objects are similar to the degree that they have features in common and do not have distinctive features. But not all comparisons produce a measure of similarity. (Consider a computer comparing two bit patterns to decide its next operation). Rather, similarity employs comparison to assess how much the entities depart from identity. Beyond the notions of a comparison and measure of departure from identity, however, similarity is open to competing interpretations, interpretations that ascribe it explanatory roles that differ widely in their strength. Here's a sample of possible positions along this continuum:

2.1. *Strong similarity*

At one extreme is the 'original sim' view that similarity is primitive in perception and cognition. Although our *judgments* of similarity might be subject to biases, there is a raw feeling of similarity that we have when we confront two objects that is fixed once and for all by our cognitive system. We'd expect similarity of this stripe to be relatively automatic, fast, perceptual, and impenetrable (i.e. unaffected by a person's other beliefs). On this view, similarity is loaded with explanatory power, for similarity relations are fundamental.

2.2. *Weak similarity*

A less extreme view holds that similarity is not fixed and irreducible, but the algorithm that computes similarity is. According to this view, similarity is a result of a well-defined, highly specialized computation. However, the computation takes place over some more-or-less open and variable set of properties of the to-be-compared entities. Similarity can be context-dependent on this view because the representations that the process uses as input can change. This view gives less explanatory force to similarity than the strong view because it demands analysis of the repre-

sentations of the objects whose similarity it computes. But similarity remains a force, for people calculate it in a nonarbitrary fashion and can use it explicitly in performing cognitive tasks, such as reasoning and categorization.

2.3. *Feeble similarity*

Another view holds that similarity is just a short-hand way of talking about property-based cognitive processing. Many kinds of cognitive functions, including those just mentioned, depend on comparison. For example, the system may need to know whether one representation matches another in order to determine whether to perform some further task. There is no need to assume, however, that the comparison processes all use the same fixed algorithm. Instead, they may have only a rough family resemblance because they are all sensitive to the extent of common and distinctive properties between entities. In other respects, the comparison processes might be different and hence some of the seeming utility of similarity may evaporate.

2.4. *No-similarity*

Finally, at the other end of the scale, is the view that similarity has negligible explanatory power. On this view, cognition operates according to some other set of principles in which similarity has no special place. We always need a way to determine identity—in order to ensure continuity in our relations with things—but similarity is not involved in such determinations. Instead, the direction of explanation is the other way round. Similarity is itself parasitic on identity—for example, identity of common features. We can make similarity judgments, but those judgments, on this view, are so context-dependent, so complex, and so disparate that understanding them would also require understanding whatever we are trying to use similarity to explain.

3. Rules as an alternative to similarity

In specifying the role of similarity in cognition, we should resolve where we stand on the continuum above. But we should also specify what the alternatives are to similarity, and this is an equally open question. We will continue to use *rules* as an umbrella term for possible alternatives, because that term seems to be the most general one capturing the properties of interest. Rules come in a variety of forms (e.g. natural language rules, logical rules) and in a variety of types (e.g. normative rules, descriptive rules, imperatives, formal rules). Our use of *rule* in this article, however, focuses on explicit mental procedures, and when we use the term without a qualifier (*linguistic, logical, etc.*), we mean specifically mental rules. Rules take one or more mental representations as input, carry out a finite number of internal steps, and produce one or more representations as output.

This broad use of *rule* is, of course, not incompatible with similarity-based pro-

cesses, since rules could be sensitive to similarity or could even compute similarity (e.g. by counting the number of properties two objects share). In our initial example, a rule could determine the similarity between the new painting and Braque's. However, rules are not limited to reacting to similarity or to assessing it, and when we talk about rules in what follows what we have in mind are mental procedures that do *not* make essential use of similarity. We're about to see some examples of rules of this sort.

One attribute that many systems employing these rules share is *compositionality*. Rule systems often include a method for deploying rules to build complex representations out of simpler components. Linguistic rules, for example, can build complex sentences, such as (1a), from simpler ones, such as (1b) and (1c):

1. a. Fred left for Chicago, and Ginger left for DC.
2. b. Fred left for Chicago.
3. c. Ginger left for DC.

Because this process can repeat without limit, the rules can form a potentially infinite number of new complex sentences from a finite vocabulary. Much the same is true of logical rules. For example, the logical rule of *And Introduction* stipulates that a conjoined sentence, such as (1a) follows deductively from its components, (1b) and (1c), taken together. Since this rule, too, can apply to its own output, it can produce an infinite number of new logical entailments (albeit dull ones, such as *Fred left for Chicago and (Fred left for Chicago and (... and (Fred left for Chicago and Ginger left for DC)...))*). Linguistic and logic rules provide a straightforward way to describe and to implement these principles of combination. And in a psychological context, rules can help explain how people perform similar tasks like constructing an interpretation of a complex sentence from interpretations of its parts or drawing deductive inferences from given information. Analogous principles might even apply beyond language comprehension and deduction to all thinking. Rules might produce all complex mental representations by combining simpler ones.

By rules, we mean to encompass many procedures for determining the validity of inferences, including 'mental models' as described by Johnson-Laird and Byrne (1991). Although these theorists avoid the use of standard logical rules, such as *And Introduction*, they do not describe an approach that is outside the boundaries of the more general rule-based framework that we outlined above. Their proposal consists of specific procedures that combine individual mental representations to form new ones (Rips, 1994).

Fodor and Pylyshyn (1988) argue that the twin representational virtues of rule-based systems are *productivity* and *systematicity*. Rule-based systems are productive in that they can produce a theoretically unlimited number of new representations, as we have just seen. They are systematic in that their ability to produce one representation implies their ability to produce certain others. If a rule can combine representations A and B, and B' is of the same type as B, then the rule can combine A with B'. For example, the rule that combines (1b) and (1c) to produce (1a) could also combine (1b) with the sentence *Joan left for Spokane* to produce *Fred left for Chicago and Joan left for Spokane*. Human thought seems to have this character as

well. Anyone who can imagine a person writing a sonnet also has the competence to imagine a person writing a ballad.

The properties of productivity and systematicity derive from the internal structure of rules—their abstract logical form and their use of variables. Productivity arises because rules can be applied in indefinitely long sequences, though they themselves are finite. For example, a rule like *Add 1 to integer N to produce the integer that is the successor of N* can be applied an infinite number of times to generate the entire class of integers. This is possible because of the presence of the variable *N* that allows the rule to apply to any integer. Systematicity arises because the abstract logical form of rules ensures that they apply to an entire class and not just to an individual object. Models of similarity generally compare representations without regard for logical form or variables. In this sense, similarity-based models seem to provide a natural contrast to rule-based ones.

4. The systems as computational devices

What kind of computational considerations can be brought to bear on the rule vs. similarity debate? Answering this question requires us to be more specific about the two kinds of systems. One fairly general approach to rules is to conceive of them as part of a *production system* (Newell, 1990). A production system consists of a large set of condition-action pairs (*production rules*) that specify an action to be performed if the condition holds. For instance, a production rule for producing the conjoined sentence in (1a) might be: IF S_1 and S_2 are each grammatical English sentences, THEN assert that S_1 and S_2 is a grammatical English sentence. You can use sets of production rules to implement most existing rule systems.

An inclusive approach to similarity-based systems is to identify them with any system whose processing of a representation is governed by the location of that representation relative to others in a continuous space. One large class of computational systems of this kind are called *dynamical systems*. These systems operate by associating representations with desired outputs, outputs that sometimes represent actions that the system can perform. It can be shown mathematically that production systems and dynamical systems have comparable degrees of computational power. Hence, on these assumptions, computational power is not sufficient to allow us to decide whether similarity or rules are preferable as a general ‘language’ of mental processing.

If power alone can’t distinguish the approaches computationally, we must rely on criteria like efficiency, robustness, flexibility, and simplicity to evaluate their computational value. Here, systems of both kinds offer advantages. Debate over their relative merits is now common in the AI literature; some of it is reviewed in Chater and Hahn’s paper. Note that in their paper, and in contrast with our position, Chater and Hahn argue that similarity should not be equated with an arbitrary associative system, in particular that connectionist systems are not similarity-based. The brunt of their effort is to specify the meaning of *rule-based* and *similarity-based* in an empirically consequential way.

Along with the virtue of compositionality, rule-based systems are often easier to analyze. We can sometimes prove that certain properties hold of rule-based systems because the rules come in discrete, explicit, and finite pieces. More generally, the ability to translate many theories directly into the kind of rules that conventional computer programs use makes many analyses easier and sometimes makes analyses possible. Moreover, the structure that rules share with natural discourse makes it easier to construct comprehensible explanations out of patterns of rule applications and to design interfaces between rule-based systems and humans.

In contrast, similarity-based models tend to be more robust to error and hardware failure because the models can more easily distribute computation across large portions of the system. They also tend to be flexible. Because such systems can generalize automatically on the basis of similarity, they can respond to stimuli that their designers didn't consider. When a computation requires retrieving information from a large data base, such systems are often efficient because they can retrieve instances in parallel that are similar to the memory probe.

Similarity-based models are closely related to conditional probability (Golden, 1988). The similarities among a set of objects, for example, maps into the probability that one of them has some property given that the others do. Prominent psychological theories have taken advantage of this relation, using similarity to predict the conditional probability of a response (e.g. Shepard, 1987), an inductive inference (Osherson et al., 1990; Sloman, 1993), or a categorization decision (e.g. Ashby, 1992).

In sum, computational considerations suggest that each approach has virtues and that these virtues are complementary. Roughly, rules provide precision, expressiveness, and generativity, and similarity provides flexibility and the means to deal with uncertainty.

5. Review of the data

We now briefly review some studies of human thinking with an eye toward answering two questions: Can we find evidence for or against rules and similarity? And, does this evidence throw further light on how to conceive of similarity? In a sense, this review wrote itself, for the distinction between similarity and rules turns up in all areas of cognitive science that deal with abstract thought.

5.1. Reasoning

The case for rules is clearest in the study of deductive inference—reasoning to certain conclusions. Several rule-based general theories of deduction have appeared that account for people's judgments of when a conclusion follows with certainty from a set of premises and when it does not (e.g. Braine et al., 1984; Rips, 1994). We know of no theories based on similarity that are successful over the same domain

(predicate logic). The possibility remains, however, that people handle certain deductively valid arguments through procedures that do not themselves grind out valid conclusions (e.g. Sloman, 1997). If so, this opens the door for similarity-based theories.

By contrast, the field of inductive inference—reasoning to uncertain conclusions—is a battleground for the rules vs. similarity issue. The evidence is compelling that people sometimes use rules to make inductive inferences. Smith et al. (1992) make the case that people explicitly use certain probabilistic rules, although their discussion is limited to a small number of them. They found evidence that people apply the statistical law of large numbers when making inductive inferences. However, Kahneman and Tversky (1972) and Tversky and Kahneman (1973) demonstrated long ago the central role of heuristics in reasoning about probability—heuristics that at best approximate normative, rule-based procedures. Similarity-based machinery could easily implement two of these heuristics: the *representativeness* heuristic (which states that people make judgments of probability and frequency by determining the similarity of a description to the category being judged or to the outcome they would expect from the process being judged) and the *availability* heuristic (which states that people judge the likelihood of a class according to the ease with which instances of the class come to mind). The availability heuristic is similarity-based as it depends on a memory retrieval process that in turn depends on the similarity between an encoded event and retrieval cues.

Osherson et al. (1986) discuss the virtues and shortcomings of seven theories of inductive inference. Four of the theories derive from logic; another assumes the construction of causal schema; and another is based on a scale of ‘conditional plausibility’. Osherson et al. (1990) conclude that each theory is likely to explain the source of some beliefs, but they also expose the shortcomings of each. A final theory is similarity-based. This theory differs from the others in applying to only a limited domain of arguments. It applies to arguments consisting entirely of category-predicate statements (e.g., *Turnips have vitamin T; therefore, parsnips have vitamin T*) that vary in the similarity among their categories or among their predicates. Further work (Osherson et al., 1990, 1994; Sloman, 1993) has developed this similarity-based approach in more detail and with a reasonable degree of empirical success. However, the framework’s domain of application remains narrow.

The evidence so far suggests that both rules and similarity have places in a complete theory of inductive reasoning. Sloman (1996) has attempted to characterize their respective roles in the course of arguing that reasoning involves two complementary systems, a rule-based one and an associative one. That characterization remains underdetermined however. In this issue, Goldstone and Barsalou argue that both kinds of processing—indeed, all aspects of cognition—have a perceptual basis. They attempt to bridge the rule/similarity divide by constructing a continuum from perception to conception in which even the most abstract representations have some perceptual basis (and in which tasks differ in how much they depend on raw perceptual input).

5.2. Categorization

The standard lore in cognitive science is that categories are composed of similar instances. However, researchers have uncovered several dissociations between categorization and similarity. Barsalou (1985) demonstrated the existence of goal-derived categories (e.g. birthday presents) which group highly dissimilar items because they serve a common goal. Rips (1989) has shown that certain properties of natural kinds can differentially affect classification and similarity judgments. In particular, similarity judgments are more sensitive to perceptual properties than to other more central properties. Central properties of animals would include such things as their genetic structure and parentage. Keil (1989) has shown that the influence of these central properties grows over the course of development. Rips and Collins (1993) have used categories whose instances are distributed on a single dimension to show that similarity judgments depend on the distance of an instance from the distribution's center; categorization, however, depends on the relative frequency of instances. Finally, Rips (1989) has shown a dissociation between similarity and categorization produced by objects with a dimension value close to those of members of a category but nevertheless outside the category's boundary (e.g. an object that looks like a US quarter but has a 3-inch diameter). Subjects judge such objects similar to the category but nevertheless excluded from it. The quadrilateral at the right of Fig. 1 provides another example of this sort.

Some of these dissociations between categorization and similarity may depend on subjects adopting an analytical mode of processing rather than responding on an immediate, intuitive basis (Smith and Sloman, 1994). Nevertheless, the variety of dissociations between categorization and similarity have led a number of researchers to reject similarity as a basis for categorization in favor of the view that categorization depends on naïve domain theories (Carey, 1985; Keil, 1989). These researchers share the belief that instances are bound into a category by virtue of a common explanatory relation (Rips, 1989). Instances are classed together because their origin or function explains them in the same way. This view of categorization generally contrasts with prototype, exemplar, and connectionist models, all of which assume that similarity is the critical category-forming relation.

Taken together, we agree that the data cited demonstrate convincingly that categorization cannot be reduced in toto to a context-independent scale of similarity. Even so, we still find those who support the view that similarity plays a substantive role in categorizing. For example, Goldstone (1994a) argues that similarity is not so context-sensitive as to rob it of explanatory power. Abundant evidence shows that judgments of similarity can accurately predict classification performance for simple stimuli (e.g. Nosofsky, 1984).

The thinking of Quine (1977) about natural kinds led him to the conclusion that science often classes entities together because of a common explanatory link, and not because of their similarity to each other. The constituents of (say) the human immune system are in no sense similar to one another; they are members of a class only in that they participate together in a specialized biological function. Their commonality derives from the scientific account of the immune system. But such

accounts—explanations—are not always available. In such cases, people may have to rely on their ‘animal sense of similarity’ by categorizing without an explanatory foundation (e.g. Smith et al., 1974). In general, people may have access to two distinct types of categorization processes—a rule-based one that is able to construct explanations, as well as a similarity-based one (Allen and Brooks, 1991; Smith and Sloman, 1994). Smith, Patalano, and Jonides in this issue report neuroimaging data showing that different parts of the brain are at work when people categorize on the basis of similarity to a stored exemplar than when they categorize according to a rule.

Keil, Smith, Simons, and Levin take issue with the Quinean hybrid model, according to which the explanatory process kicks in only after the stage has been set by similarity. They argue that, from the start, children struggle to explain the internal structure of the objects they encounter. Similarity-based processing, they argue, depends as much on explanation as explanation does on a stable similarity space. Hampton, in contrast, argues for the primacy of similarity in categorization. He shows strong correlations between people’s judgment of the typicality of an item in a category, which he takes as a measure of similarity to a prototype, and the probability that people will classify the item as a member of the category. The correlations are not perfect, but Hampton is able to explain away most of the discrepancies by appealing to different demands of the typicality and categorization tasks.

5.3. *Metaphor and analogy*

One view of metaphor is that it is not substantially different from analogy. For example, a metaphor, such as (2), depends on the similarity between the metaphor’s *topic*, Mozart, and its *vehicle*, pastry chef:

(2) Mozart is the pastry chef of composers.

Such an explicit comparison process is part of both the salience imbalance model of metaphor (Ortony, 1979) and the structure mapping approach to analogy (Gentner, 1983). On this view, understanding (2) requires comparing Mozart to pastry chefs in order to find a distinctive property of pastry chefs that can be ascribed to Mozart (perhaps the ability to create elaborate, beautiful, and enjoyable objects).

One provocative reaction to this view has been to deny that metaphor involves a similarity comparison. *Literal* statements that have the form *An X is a Y* often mean that *Xs* are included among *Ys* (e.g. *An anteater is a mammal* means that anteaters are in the class of mammals). Glucksberg and Keysar (1993) argue that metaphors of this form should be understood as category inclusion statements too: The metaphor *An X is a Y* means that *Xs* belongs to the category that *Y* typifies, not that *X* and *Y* are similar. For example, (2) means that Mozart belongs to the class of people who create elaborate, beautiful, and enjoyable objects. This debate about the interpretation of metaphor thus turns in part on the role of similarity in categorization. If categorization reduces to similarity, the difference between the two positions is

negligible. But if Glucksberg and Keysar are right that metaphor comprehension is categorization and if metaphor comprehension depends on those aspects of categorization that are not reducible to similarity, then rules may govern metaphor in the same way they govern categorization.

According to both theories, understanding metaphors and analogies requires assigning properties of one entity to another, as our example suggests. So one framing of the basic issue in this literature is the extent to which such property-mapping emerges directly from alignment and comparison as opposed to rule-governed analysis and synthesis. In their paper, Gentner and Medina, make the case for alignment and comparison, not just in analogy, but as central processes in learning and in the abstraction and application of all kinds of rules.

5.4. *Decision-making*

Two main theories have dominated the study of decision-making. On one hand, there are models that mathematically combine subjective probability (e.g. your belief about the likelihood of rain) and subjective value (e.g. the strength of your desire to stay dry) to produce subjective expected utility (the overall desirability of taking your umbrella). (See Savage, 1954; cf. Fishburn, 1988). On the other hand, there are models that rely on heuristics to explain how people behave when confronted with choices (Kahneman et al., 1982). One of these heuristics, representativeness, claims that people make judgments of probability or frequency by determining the similarity of a description to a category or to the outcome of a process, as we mentioned earlier. Medin et al. (1995) have elaborated this observation by pointing out a number of parallels between choice and similarity. They suggest the possibility that choice may reflect the similarity of options to an ideal option. Rules may be staging a comeback in this area, however, in the guise of theories that use argumentation or explanation to guide choice (e.g. Shafir et al., 1993; Hogarth and Kunreuther, 1995).

5.5. *Similarity judgments*

Similarity can be regarded either as a form of judgment—what people say when asked to compare two entities—or as a theoretical comparison process that generates a similarity measure. Up to now, we have been considering similarity in the second sense, but we should also look at the way people judge similarity overtly. The need to distinguish similarity as a task from similarity as a theoretical construct arises because the study of the task once again raises the similarity versus rules issue. Sometimes automatic comparison processes seem to dominate similarity judgments. For example, Garner (1974; see also Goldstone, 1994b) has shown that when people must discriminate visual patterns according to their values on one dimension, variation on an irrelevant dimension can slow them down. Featural similarities on the second dimension are inhibitory even though they are irrelevant. In contrast, similarity judgments sometimes reflect the outcome of selective analysis. For example, falcons are judged more similar to chickens when the set of judged items includes

wasps than when it includes sparrows (Tversky, 1977). The context in which judgments are made can selectively activate properties (for an extreme example of a ‘cognitively penetrable’ similarity task, see Bassock and Medin, 1997). Selective analysis may or may not implicate rules; nevertheless, different similarity tasks clearly require different explanatory models.

6. Conclusion

By characterizing similarity as we have, we believe that we have eliminated the possibility that all thought is similarity-based. Some rule-based mechanisms are necessary to explain (a) our competence to use language systematically and productively (Fodor and Pylyshyn, 1988); (b) the results of Smith et al. (1992); and (c) the sense of certainty associated with some inferences even in unfamiliar domains. To elaborate on (c), consider the argument *A, B, C, and D, therefore A*. We believe that all reasonable people consider this a conclusive argument. The certainty associated with it can be understood as a by-product of the application of a deductive inference rule like AND Elimination. It cannot be attributed to similarity-based computation because the argument *A, B, C, and D, therefore A, B, C, D, and E* is not conclusive, even though the premise is more similar to the conclusion than in the previous argument.

Of course, a dual view remains viable—that people are capable of explicit rule application, but they also have and use a special-purpose similarity-based inference system (Sloman, 1996). On this view, similarity-based processing obeys fundamentally different principles and uses different mechanisms than does rule-based processing.

How should we conceive of similarity? Which of the views outlined earlier should we accept (if any)? We believe that we can reject the Strong Similarity view, at least in the domain of abstract concepts. Similarity is too manipulable and too complex to be considered an irreducible relation. This may not be true in the perceptual domain, though. Analyses of color and of musical pitch have yielded similarity structures that may capture regularities across contexts (Shepard, 1980). No such reliable structure has been discovered at more abstract levels of analysis. Moreover, for similarity to have one of the virtues that we ascribe it—flexibility—it can’t represent a set of fixed relations. Context-dependency must be intrinsic to its calculations.

Perhaps similarity can maintain some explanatory credentials, however, provided that it can be fitted with proper restraints (Goldstone, 1994a). In the Weak View (maybe even in the Feeble View), similarity remains a key building block in some forms of cognition. By giving it some special role, we can buy a lot of computational power while explaining why we often display a sensitivity to similarity, even when confronted simultaneously by a conflicting, more justifiable conclusion (Sloman, 1996). For unidimensional perceptual judgments, especially, similarity is hard to eliminate. Although it is possible to explain similarity away even in this context, attempts to do so seem to us forced and artificial.

Just what special role similarity has is yet to be decided. Rules are powerful entities, and the jury has yet to decide whether the data can support an autonomous notion of similarity. The purpose of this special issue is to help arrive at a more informed conclusion.

Acknowledgements

Steven Sloman was funded by NIMH grant MH51271 to Barbara Malt and Steven Sloman; Lance Rips by NSF grant SBR-9514491.

References

- Allen, S.W., Brooks, L.R., 1991. Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General* 120, 3–19.
- Ashby, F.G., 1992. Multidimensional models of categorization. In: Ashby, F.G. (Ed.), *Multidimensional Models of Perception and Cognition*. Erlbaum, Hillsdale, NJ.
- Barsalou, L.W., 1985. Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning Memory, and Cognition* 11, 629–654.
- Bassock, M., Medin, D.L., 1997. Birds of a feather flock together: similarity judgments with semantically-rich stimuli. *Journal of Memory and Language* 36, 311–336.
- Braine, M.D.S., Reiser, B.J., Rumin, B., 1984. Some empirical justification for a theory of natural propositional logic. In: Bower, G.H. (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Thinking*, Vol. 18. Academic Press, New York, pp. 313–371.
- Carey, S., 1985. *Conceptual Change in Childhood*. MIT Press, Cambridge, MA.
- Fishburn, P.C., 1988. *Nonlinear Preference and Utility Theory*. Johns Hopkins Press, Baltimore, MD.
- Fodor, J.A., Pylyshyn, Z.W., 1988. Connectionism and cognitive architecture: a critical analysis. *Cognition* 28, 3–71.
- Garner, W.R., 1974. *The Processing of Information and Structure*. Erlbaum, Hillsdale, NJ.
- Gentner, D., 1983. Structure-mapping: a theoretical framework for analogy. *Cognitive Science* 7, 155–170.
- Glucksberg, S., Keysar, B., 1993. How metaphors work. In: Ortony, A. (Ed.), *Metaphor and Thought*, 2nd edition. Cambridge University Press, Cambridge.
- Golden, R.M., 1988. A unified framework for connectionist systems. *Biological Cybernetics* 59, 109–120.
- Goldstone, R.L., 1994a. The role of similarity in categorization: providing a groundwork. *Cognition* 52, 125–157.
- Goldstone, R.L., 1994b. Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General* 123, 178–200.
- Goldstone, R.L., Medin, D.L., Gentner, D., 1991. Relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology* 23, 222–262.
- Goodman, N., 1955. *Fact, Fiction, and Forecast*. Harvard University Press, Cambridge.
- Hogarth, R.M., Kunreuther, H., 1995. Decision making under ignorance: arguing with yourself. *Journal of Risk and Uncertainty* 10, 15–36.
- Johnson-Laird, P.N., Byrne, R.M.J., 1991. *Deduction*. Erlbaum, Hillsdale NJ.
- Kahneman, D., Tversky, A., 1972. Subjective probability: a judgment of representativeness. *Cognitive Psychology* 3, 430–454.
- Kahneman, D., Slovic, P., Tversky, A., 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge.
- Keil, F.C., 1989. *Concepts, Kinds and Cognitive Development*. MIT Press, Cambridge, MA.

- Nosofsky, R.M., 1984. Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning Memory, and Cognition* 10, 104–114.
- Medin, D.L., Goldstone, R.L., Markman, A.B., 1995. Comparison and choice: relations between similarity processes and decision processes. *Psychonomics Bulletin and Review* 2, 1–19.
- Murphy, G.L., Medin, D.L., 1985. The role of theories in conceptual coherence. *Psychological Review* 92, 289–316.
- Newell, A., 1990. *Unified Theories of Cognition*. Harvard University Press, Cambridge, MA.
- Ortony, A., 1979. Beyond literal similarity. *Psychological Review* 86, 161–180.
- Osherson, D., Smith, E.E., Shafir, E., 1986. Some origins of belief. *Cognition* 24, 197–224.
- Osherson, D., Smith, E.E., Meyers, T.S., Shafir, E., Stob, M., 1994. Extrapolating human probability judgment. *Theory and Decision* 36, 103–129.
- Osherson, D., Smith, E.E., Wilkie, O., Lopez, A., Shafir, E., 1990. Category-based induction. *Psychological Review* 97, 185–200.
- Quine, W.V., 1977. Natural kinds. In: Schwartz, S.P. (Ed.), *Naming, Necessity, and Natural Kinds*. Cornell University Press, Ithaca, NY.
- Rips, L.J., 1989. Similarity, typicality, and categorization. In: Vosniadou, S., Ortony, A. (Eds.) *Similarity and Analogical Reasoning*. Cambridge University Press, Cambridge.
- Rips, L.J., 1994. *The Psychology of Proof: Deductive Reasoning in Human Thinking*. The MIT Press, Cambridge, MA.
- Rips, L.J., Collins, A., 1993. Categories and resemblance. *Journal of Experimental Psychology: General* 122, 468–486.
- Savage, L.J., 1954. *The Foundations of Statistics*. Wiley, New York.
- Shafir, E., Simonson, I., Tversky, A., 1993. Reason-based choice. *Cognition* 49, 11–36.
- Shepard, R.N., 1980. Multidimensional scaling, tree-fitting, and clustering. *Science* 210, 390–398.
- Shepard, R.N., 1987. Towards a universal law of generalization for psychological science. *Science* 237, 1317–1323.
- Sloman, S.A., 1993. Feature-based induction. *Cognitive Psychology* 25, 231–280.
- Sloman, S.A., 1996. The empirical case for two systems of reasoning. *Psychological Bulletin* 119, 3–22.
- Sloman, S.A., 1997. Categorical inference is not a tree: the myth of inheritance hierarchies. *Cognitive Psychology*.
- Smith, E.E., Langston, C., Nisbett, R., 1992. The case for rules in reasoning. *Cognitive Science* 16, 1–40.
- Smith, E.E., Shoben, E.J., Rips, L., 1974. Structure and process in semantic memory: a featural model for semantic decisions. *Psychological Review* 81, 214–241.
- Smith, E.E., Sloman, S.A., 1994. Similarity– vs. rule–based categorization. *Memory and Cognition* 22, 377–386.
- Tienson, J.L. Jr., 1988. Resemblance and general terms. *Philosophical Studies* 54, 87–108.
- Tversky, A., 1977. Features of similarity. *Psychological Review* 84, 327–352.
- Tversky, A., Kahneman, D., 1973. Availability: a heuristic for judging frequency and probability. *Cognitive Psychology* 5, 207–232.