As I see it, similarity is a metric, whereas rules are computational operations. Rules (or what I have called "algebraic" operations [Marcus 2001]) can be used to construct an algorithm that computes the similarity between two entities (e.g., by calculating the cosine between two vectors or by tabulating what proportion of features are shared by two members of a category), or to construct something entirely different (e.g., to borrow an example from the late Stephen Jay Gould, checkbook balancing). Although many algorithms for computing similarity depend on the application of rules, not all rules compute similarity, and – to the extent that evaluations of similarity might come about implicitly as the product of memory retrieval rather than explicitly and algorithmically – some computations of similarity may not depend on rule-application. Most of the tests that purport to distinguish rules from similarity are really about something else, such as a vaguely kindred distinction between rules and memory storage, that Pinker, Clahsen, and myself have argued for elsewhere (Clahsen 1999; Marcus 2001; Pinker 1999). Our argument has been that the past tense of an English verb (for instance), may be produced by one of two pathways: a system that applies a concatenation operation (rule) to a verb stem, or an alternative system that retrieves (irregular and in some cases regular) past-tense forms from an associative memory. That memory system is sensitive to similarity; such similarity might in principle be computed via a rule (as it has in some computational models of irregular inflection). All of this may be merely terminological, but if one adopts the perspective I have developed herein, it follows that tests for discerning rule-application and similarity should not try to neatly categorize all computations into one of two mutually exclusive bins (Rules vs. Similarity, not both) but, rather, seek to answer two separable questions: (a) whether a particular given computation is instantiated by means of the application of algebraic operations (such as the instructions in a microprocessor), and (b) whether that computation depends on calculating (perhaps as intermediate values) the extent to which two inputs share representational material. To see why the two questions are to some extent dissociable, consider that a simple web search engine that returned a value based on the proportion of matching words out of total words would be both rule-based and similarity-computing; an alphabetizing (sorting) algorithm, in contrast, might be implemented via rule but leave no room for similarity. A system that relied purely on the strengths of particular memory traces might be an instance of similarity-based systems that did not invoke rules. The putative rules-and-similarity contrast fails not because they are opposites, but because they are orthogonal, answering separate questions, one about the nature of the operations with which a given algorithm is implemented, the other about what that algorithm computes. It is with good reason that efforts to pigeonhole them in separate bins have failed.

ACKNOWLEDGMENTS

This commentary is supported by grants from the NIH and HFSP.

Digging beneath Rules and Similarity

Arthur B. Markman, Sergey Blok, Kyungil Kim, Levi Larkey, Lisa R. Narvaez, C. Hunt Stilwell, and Eric Taylor

Department of Psychology, University of Texas, Austin, TX 78712.
markman@psy.utexas.edu **blok@psy.utexas.edu**
kyungil@psy.utexas.edu **larkey@mail.utexas.edu**
grimmlr@mail.utexas.edu **stilwell@psy.utexas.edu**
eric@gregalo.com
http://www.psy.utexas.edu/psy/faculty/markman/index.html

Abstract: Pothos suggests dispensing with the distinction between rules and similarity, without defining what is meant by either term. We agree that there are problems with the distinction between rules and similarity, but believe these will be solved only by exploring the representations and processes underlying cases purported to involve rules and similarity.

Pothos suggests that the distinction between rules and similarity is not necessary, because similarity processes alone can be used to explain data previously thought to require rules. We agree that the distinction between rules and similarity is problematic, but disagree that the solution is to eliminate the problem by focusing selectively on similarity. Instead, it is more important to focus on underlying representational and processing issues related to cognition. Ultimately, it may be possible to use rules or similarity (or both) to model these cognitive processes. However, progress on understanding cognition will require focusing on a finer-grained set of issues than a distinction between rules and similarity permits. To illustrate this point, we first discuss three dimensions that are often correlated with the rules versus similarity distinction. We demonstrate that the poles of these dimensions can be implemented using either rule-based or similarity-based models.

Three dimensions of processing. Intuitively, rules and similarity appear to differ along (at least) three different dimensions. Rules are assumed to be abstract, while similarity is assumed to be based in the content of the domains compared. Rules pre-compile the relative salience of properties, whereas similarity determines the salience of properties contextually. Rules are involved in the application of knowledge, while similarity is involved in the transfer of knowledge to new domains.

Degree of abstraction. Similarity comparisons are assumed to be relatively concrete and to preserve specific aspects of the items compared. In contrast, rules are assumed to capture abstract regularities across situations. Figure 1a illustrates this point by using bolded boxes to indicate the typical way that similarity and rules are linked to degrees of abstraction.

There is no requirement that similarity models involve concrete aspects of items, nor is it required that rules focus on abstractions. For example, structural approaches to analogy and similarity (Gentner & Markman 1997; Hummel & Holyoak 1997) permit the outcome of a comparison to be an abstract relational match between situations. Likewise, rule-based models of automaticity assume that people may form specific productions that fire when an automatized sequence should be carried out (Anderson 1983). For example, Anderson (1983) suggests one might have a production that fires only when the goal is to dial a particular friend's phone number.

The influence of general and specific information on cognitive processing is an important area of research, and one that deserves much further study (Medin & Ross 1989). However, as illustrated in Figure 1, it is possible to implement models of processes requiring various degrees of representational specificity using either rules or similarity.

Determination of salience. Models incorporating rules and similarity often differ in the expectation that the importance of representational elements will be determined contextually (Fig. 1b). Models of similarity assume that contextual factors affect salience (Gati & Tversky 1984; Gentner 1983; Medinet al. 1993). For example, Gentner's (1983) systematicity principle suggests that a given property is more important for a similarity comparison when it is relationally connected to other matching information than when it is not. Rule-based models typically assume that only salient properties are encoded in the rules as part of the process that extracts rules during learning. Once again, however, there is no principled reason for this relationship between rules and similarity. Most similarity models, while highlighting the importance of contextually derived factors in the calculation of similarity, also assume that particular features may have context-neutral degrees of importance (Tversky 1977). Furthermore, production-system models often have a limited-capacity working memory that determines both the information that can be matched to the preconditions of productions as well as the strength those properties will have in matching against rules in memory (e.g., Anderson et al. 1996). Thus, processes that influence the salience of properties are critical for understanding cognitive processing, but they do not strongly favor similarity models over rule-based models.

Knowledge application. A third distinction is between the appli-

https://doi.org/10.1017/S0140525X05390017 Published online by Cambridge University Press

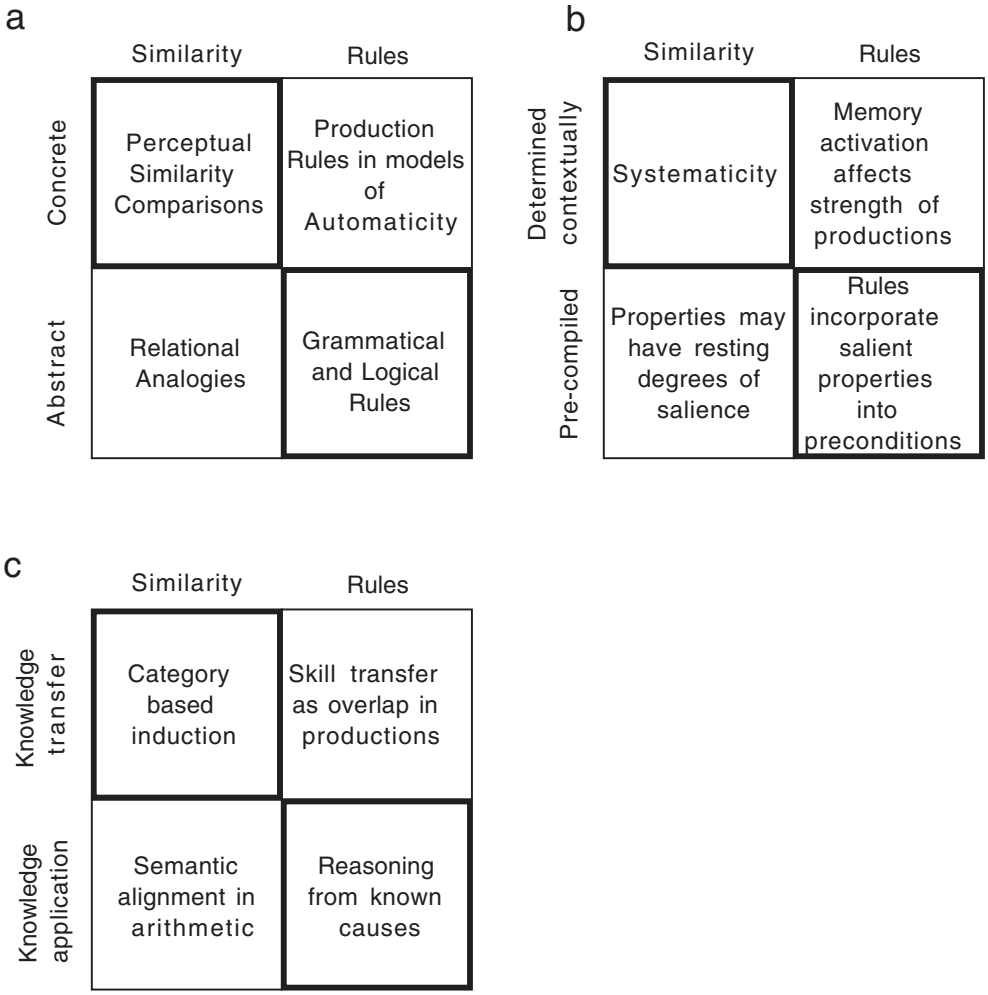


Figure 1 (Markman et al.). Three dimensions of cognitive processing typically thought to be correlated with the distinction between similarity and rules. As these panels show, the poles of these dimensions need not reflect underlying use of similarity or rules.

cation and transfer of knowledge. Models of similarity are often applied in contexts in which knowledge from one domain is being transferred to another. For example, Osherson et al.'s (1990) model of category-based induction assumes that the strength of belief that a property is true of some new category depends strongly on the similarity of that category to categories known to have that property. In contrast, models of knowledge application often involve rules. For example, models assuming that people have domain theories, assume that causal relations are represented as rules that can be applied to new causal situations in the same domain (e.g., Kim & Ahn 2002).

As shown in Figure 1c, models of knowledge application may also be similarity-based. For example, Bassok et al. (1998) have shown that arithmetic word problems written by undergraduates (who should have a strong domain-general understanding of the rules of arithmetic) are strongly constrained by the semantic content of the problems. These results suggest that undergraduates are comparing the structure of the mathematics domain to the structure of the semantic domain during knowledge application.

Finally, knowledge transfer need not be modeled by using similarity comparisons. For example, Singley and Anderson (1989) provide an extended demonstration that transfer between cognitive domains can be conceptualized as occurring when performance in each domain requires some of the same production rules.

Conclusion. The three examples in this commentary are not meant to be exhaustive. Instead, they represent three important dimensions of processing whose poles are often assumed to be associated with the distinction between similarity and rules. Pothos has pointed out cases where models of similarity can account for aspects of processing typically thought of as requiring rules. Our analysis suggests that similarity and rule-based models are both sufficiently powerful to account for the same range of phenomena. Thus, research must focus on extracting the relevant dimensions underlying cognitive processing and understanding their influence on thought. The implementation of these dimensions in models involving either similarity or rules should be of secondary importance.

https://doi.org/10.1017/S0140525X05390017 Published online by Cambridge University Press