Cornell Library Annex Interlibrary Loan

From: Olin Interlibrary Services <olin-ils-lending@cornell.edu>
Sent: Friday, December 02, 2016 1:17 PM
To: Cornell Library Annex Interlibrary Loan
Subject: ILL Request TN1320026

PLEASE SCAN THE FOLLOWING ARTICLE AND SEND A PDF TO THE SERVER. PLEASE SCAN THIS E-MAIL AT THE FRONT OF THE DOCUMENT. REPLY TO THIS E-MAIL AFTER FILLING IN THE BOXES BELOW:

__ PDF TO SERVER
__ Number of pages

COPY REQUEST - TN: 1320026

Location: Library Annex
CALL NUMBER: BF30 .P97

JOURNAL or MONOGRAPH TITLE: Psychologica Belgica.Psychologica Belgica. GOLDSTONE, RL: MAINSTREAM AND AVANT-GARDE SIMILARITY

VOL/ISSUE/DATE/PGS: 35 2-3 1995 tp,copyrt,145-165

ARTICLE: GOLDSTONE, RL: MAINSTREAM AND AVANT-GARDE SIMILARITY

ILL number: BNG 170270183
Requesting Library:
ILL Bartle Library

Request type: SCRLC 30.00IFM

NOTICE: THIS MATERIAL MAY BE PROTECTED BY COPYRIGHT LAW (TITLE 17 U.S. CODE) This copy has been provided from the collections of the Cornell University Library in conformance with U.S. copyright law. It is to become the property of the requestor, and is only to be used for personal study, scholarship, or research. Any other use may require the permission of the copyright owner. It is the responsibility of the person requesting this item to make an independent legal assessment of any proposed use and to secure any necessary permissions.

If you are unable to fill this request, please check one of the reasons below:
___In use
___In process
___Lost
___Non-circulating
___Not owned
___On order
___At bindery
___Lack volume or issue
___Not on shelf
___On reserve

PSYCHOLOGICA BELGICA

QUARTERLY JOURNAL PUBLISHED BY THE
BELGIAN PSYCHOLOGICAL SOCIETY

1995
vol. 35
nr. 2/3

PUBLIÉ AVEC LE CONCOURS DE LA COMMUNAUTÉ FRANÇAISE, SERVICE DE LA RECHERCHE SCIENTIFIQUE ET LE CONCOURS DE LA FONDATION UNIVERSITAIRE - UITGEGEVEN MET DE FINANCIËLE STEUN VAN HET MINISTERIE VAN DE VLAAMSE GEMEENSCHAP-DEPARTEMENT ONDERWIJS EN MET DE FINANCIËLE STEUN VAN DE UNIVERSITAIRE STICHTING

BELGISCHE VERENIGING VOOR PSYCHOLOGIE
SOCIÉTÉ BELGE DE PSYCHOLOGIE

Robert SCHETS
Boulevard Mettewie 79 b 43
1080 BRUXELLES

BACK VOLUMES

Each volume: BEF 1.500

Single issues: BEF 900

International Standaard Serial Number:
BE ISSN 0033-2879

LISTED IN

- *Abstracts Journal of the Institute of Scientific Information of the CIS Academy of Sciences*
- Arnold & Doyle Jr's *Education/Psychology Journals: A Scholar's Guide*
- *Bibliografie Nederlandse Sociale Wetenschappen* (R.U. Utrecht)
- *Bowker Serial Directories*
- *Current Contents: Behavioral, Social, and Management Sciences*
- *Psychological Abstracts*
- *INIST* (C.N.R.S., France)
- *Ulrich's International Periodicals Directory*

© Belgische Vereniging voor Psychologie — Société Belge de Psychologie

No parts of this book may be reproduced in any from, by print, photoprint, microfilm or any other means without written permission from the publisher.

Psychologica Belgica
1995, 35-2/3, 145-165.

MAINSTREAM AND AVANT-GARDE SIMILARITY

Robert L. GOLDSTONE
Indiana University, USA

In the first part of this article, empirical evidence is reviewed that suggests a substantial amount of flexibility and context-sensitivity in people's judgments of similarity. Three examples of flexible similarity from our laboratory are considered in detail. In the second part of the article, evidence for relatively constrained, invariant similarity assessments is considered. In the final section, a resolution to these apparently contradictory views on similarity is proposed. Assessments of similarity are used to make inferences from one entity to another. In some situations, flexible similarity is needed to tailor inferences to one's knowledge of the entities and their relations. In other situations, particularly those in which specific knowledge is missing or unavailable, a relatively constant similarity is needed to establish generally permissible inferences. Thus, the flexibility and stability of similarity may reflect its different cognitive uses.

For a long time, cognitive psychologists have attested to the fundamental cognitive importance of similarity (Attneave, 1950; Gregson, 1975; James, 1890). Assessments of similarity are often seen as the foundation for cognitive acts ranging from problem solving to categorization to decision making to memory retrieval. Given the varied uses to which similarity is put, it is likely that similarity assessments will be subserved by several processes. It is not even clear that similarity is a single thing at all. The similarity of computers to televisions is not of the same type as the similarity of computers to abacuses, and neither is like the similarity of computers to life preservers. If different similarities are themselves not similar, then one begins to worry that similarity may not be a single, coherent or useful notion at all (Goodman, 1972).

Rather than adopt this nihilistic stance toward similarity, I will argue that similarity does have multiple purposes that stretch it in different directions, but that these conflicting pressures can be reconciled. There is reason for optimism that similarity is systematic, if we know enough about the entities being compared, their context, and the perspective of the judge who is making the comparison. The conflicting pressures that I will be concerned with in this article are the pressures for flexible, highly contextualized similarity on the one hand, and well-constrained and stable similarity on the other hand. Similarity

I wish to thank Bruce Goldstone, Douglas Medin, Paula Niedenthal, Robert Nosofsky, Philippe Schyns, Richard Shiffrin, Linda Smith, and Jean-Pierre Thibaut for their useful comments and suggestions. This research was funded by National Science Foundation Grant SBR-9409232.

Correspondence concerning this article should be addressed to Robert Goldstone, Psychology Department, Indiana University, Bloomington, Indiana 47405, USA.

must be flexible enough to accommodate situational idiosyncrasies, but must also be stable enough to ground cognition. Stability without flexibility results in judgments that are not well tuned to a judge's specific demands. Flexibility without stability runs the risk of producing judgments that do not take advantage of the information that remains constant across many situations.

The plan of the paper is to first present the two faces of similarity, as flexibly tuned to situations, and as a stable, common ground. Experimental evidence from our laboratory will be reviewed that suggests strong and surprising contextual influences on similarity. At the same time, experimental and theoretical arguments suggest limits to this flexibility. The third section of the paper presents first steps toward a reconciliation between the needs for stability and flexibility in similarity.

The Flexibility of Similarity

There is an approach to similarity that would be very attractive were it also successful. This approach assumes that similarity is based on a fixed set of features or dimensions. Each entity (e.g., objects, people, words, events, pictures, or sounds) to be compared is first assigned a representation, in terms of features (Tversky, 1977), or values along continuous dimensions (Shepard, 1962). The representations are then compared for overlap or distance. Importantly, the featural or dimensional representations are determined before the comparison process takes place. This approach is attractive because it permits many formal techniques for representing similarity, such as multidimensional scaling (Shepard, 1962). If entities do not have singular representations (their coordinates in space) that can be assigned prior to the similarity computation, then multidimensional spaces that represent each entity by a point in space will not have psychological validity.

More generally, the assumption of a fixed set of features for representing compared entities is useful because it allows a limited repertoire of "building blocks" to generate a virtually infinite array of entities. The economy of combining together a small set of static elements to create a large set of representations is undeniable. One can imagine that there are separate channels for extracting various sources of information about an entity, and that these same channels can be used for any number of entities. By this line of thought, establishing that both strawberries and fire engines have the feature red can be done by the same color processing mechanism, and this mechanism can operate in ignorance of other processing. Featural and dimensional analyses, then, offer the advantages of modularity and building-block approaches to cognition, but only if entities can indeed be represented by features that do not depend upon each other or the context-at-large.

Although life would be simpler if the representation of entities was context invariant, a substantial amount of evidence suggests otherwise. A single estimate of similarity cannot be assigned to a pair of entities irrespective of their context. Of course, "Context" is a catch-all word that must be further delineated to avoid vagueness. I will consider four varieties of contextual influence, in decreasing order of their breadth. That is, as we continue, the contexts become progressively more narrowly defined.

Perspectival Context

Similarity is influenced by the comparison-maker's personal perspective or abstract viewpoint (for some clever examples, see Murphy & Spalding, this volume). This perspective may be permanent (once acquired) or transitory. Some of the best examples of relatively permanent perspectival changes to similarity come from an examination of cultural (Whorf, 1941) and novice/expert differences. For example, once a person becomes a wine expert, it is fairly difficult for them to return to the phenomenological state of somebody who cannot distinguish Bordeauxs from Burgundies. On the road to becoming an expert, the similarities that a person spontaneously notices changes. In one experiment, Sjoberg (1972) asked clothing experts (salespeople in a textiles shop) and dog experts (dog trimmers and kennel club members) to rate similarities between pairs of dogs and pairs of cloth. The experts gave lower similarity ratings to the objects from their own domain of expertise, paralleling cross-cultural studies that indicate greater differentiation between culturally important objects (Diamond, 1966).

Research in my laboratory has explored the development of expertise during an experimental session, and its influence on subsequent perceived similarities between objects (Goldstone, 1994a). In the first phase of the experiment, subjects were trained to make different categorizations. As shown in Figure 1, the stimuli were 16 squares that possessed one of four levels of size and color. Subjects were placed in one of three categorization conditions (ignoring other conditions that are not important for the current purposes): size relevant (squares with size values of 1 and 2 belonged to Category A, squares with size values of 3 and 4 belonged to Category B. Figure 1 shows the categorizations for this condition), brightness relevant (squares with brightness values of 1 and 2 belonged to Category A, while squares with brightness values of 3 and 4 belonged to Category B), and no categorization. After categorization training, subjects were instructed to make same/different judgments; two squares were displayed and subjects indicated whether they were physically identical or not. Subjects were told that the squares might vary slightly in either size or brightness. Subjects' perceptual sensitivity was calculated as a function of their ability to correctly respond "same" when presented with two identical squares,

and to correctly respond "different" when presented with two slightly different squares.

The first phase of categorization training influenced the similarity of different squares, as measured by subjects' sensitivity in distinguishing them. In particular, subjects who were originally trained to categorize on the basis of size were better able to discern that two squares differed slightly on size. Subjects who were trained to become brightness "experts" were better able to discern that squares differed on the brightness dimension. The increase in sensitivity extended across the entire trained dimension, but was greatest at the boundary between the two categories — that is, between values 2 and 3. This pattern of sensitization is reminiscent of categorical perception, in which

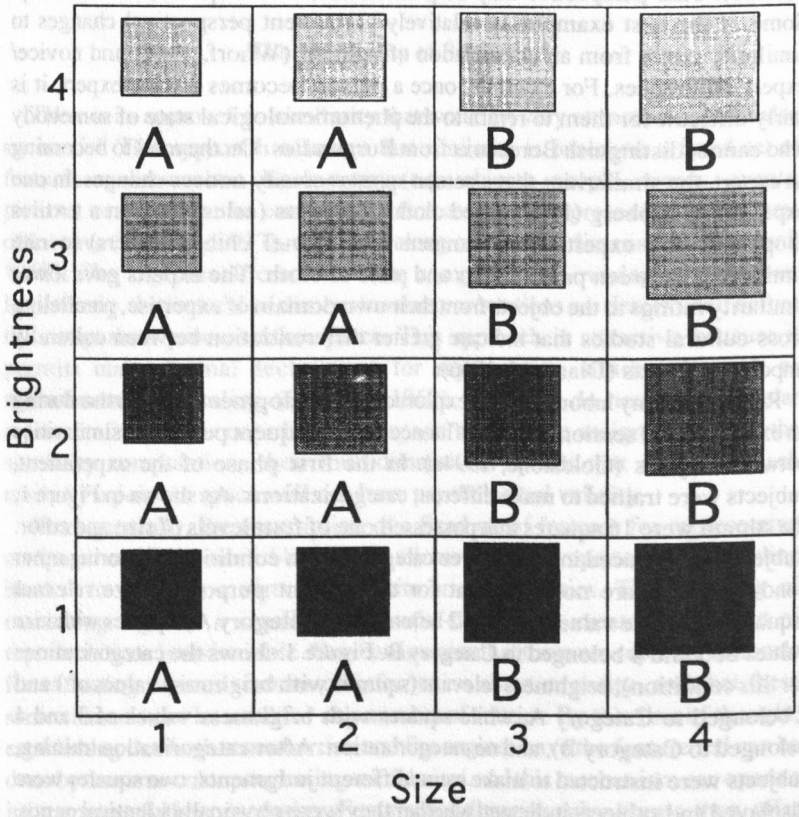

Figure 1. Sample stimuli from Goldstone (1994a). Stimuli varied on two dimensions. The letters "A" and "B" reflect the category labels for the size categorizing subjects.

subjects' ability to discriminate values on a dimension is greatest when the values belong to different categories (Harnad, 1987). Generally speaking, categorization training had the effect of increasing subjects' sensitivity, but there was one case where training decreased perceptual sensitivity. Namely, when subjects learned to categorize on the basis of brightness, their subsequent size discriminations were less accurate than were those of the control subjects who did not participate in categorization training. Thus, although the largest effect of training was to decrease the similarity (increase the discriminability) of objects that varied on a dimension that was relevant for categorization, there was also a weaker trend for training to increase the similarity of objects that varied on a dimension that was irrelevant for categorization.

This study is interesting because the influence of training seems to have been at a fairly perceptual level. Categorization training may actually change how similar objects *appear* to us. Further support for this possibility comes from a replication of the experiment that used dimensions that are typically "fused" for people (Goldstone, 1994a). The size and brightness of a square can be separately processed with ease. The brightness (related to the amount of light in the color) and saturation (related to the purity of the color) of a square are normally integrated dimensions; people have a difficult time attending to just one of the dimensions and naturally see the dimensions as comprising the single dimension of "color" (Burns & Shepp, 1988). However, when our subjects received categorization training that was based on only saturation (or brightness), then they eventually learned the proper categorization. Furthermore, after training, their discriminations of saturation differences were selectively sensitized. Training people not only can change their ability to notice dissimilarities along preexisting dimensions, but can also cause people to reorganize their perceptual world.

Recent (Laboratory) Context

The previous examples demonstrate fairly long-lasting shifts in perspective which alter impressions of similarity. Reversible and transitory shifts also occur. Contexts that are created within a laboratory experiment can temporarily alter similarities. Tversky (1977) obtained evidence for an extension effect, according to which features are more influential to similarity judgments when they vary within an entire set of stimuli (also see Sjoberg, 1972). Some of his subjects rated the similarity of pairs of South American countries, others rated the similarity of European countries, and still other subjects rated the similarity of two South American countries on some trials and two European countries on other trials. The similarity ratings from this last group were significantly higher than from the other groups, presumably because the features "South American" and "European" became important when they varied across the entire set of stimuli.

Another intuitive and clever example of how early presentations can influence later presentations comes from Kelly and Keil's (1987) study of similarity judgments across very different domains. Subjects who did not receive metaphors relating items from two domains (periodicals and food), showed less sensitivity on later tests to metaphorical similarities than subjects who did receive such metaphors. For example, subjects who received the metaphor "The New Yorker is the quiche of newspapers and magazines" gave higher similarity ratings to cross-domain pairs that had similar values on a tastefulness dimension (e.g., *steak* and *Sports Illustrated*) than did subjects who were not given these metaphors. In sum, how one comparison is interpreted depends on the comparisons that preceded it.

Concurrently Displayed Information

The similarity of entities is influenced by information which is displayed concurrently with the compared entities. Several examples of concurrent contextual information are based on the insight that an object's interpretation is altered by its surround. For example, Barsalou (1982) found that *raccoon* and *snake* were judged to be less similar when no explicit context was provided than when a context was established by placing the word *pets* above the comparison. Items that are seemingly quite different (e.g., *children* and *jewelry*) were rated as highly similar when they were placed in a context ("things to retrieve from a burning house") that highlighted a respect in which they were similar (Barsalou, 1983). Similar contextual manipulations have even produced ordinal reversals of similarities. For example, *white* is judged to be more similar to *gray* than is *black* when the context is "hair," but the opposite trend is found when the context is "clouds" (Medin and Shoben, 1988).

Asch (1952) argued that the interpretation of a feature was affected by the context defined by the other features in the object. He found that the person description {quick, skillful, helpful} was more similar to {slow, clumsy, helpful} than it was to {quick, clumsy, helpful} even though the first and third descriptions had more features in common. Asch argued that the feature *helpful* was interpreted unusually when it was combined with *quick* and *slow* (e.g., "perhaps she's clumsy because she's working too quickly, implying that she may only *appear* to be helpful."). In general, words have multiple meanings and different nuances within a single meaning, and the nuance or interpretation which is brought to mind and influences similarity changes with disambiguating context. One experiment that we conducted (Medin, Goldstone, & Gentner, 1993) provided an interesting example of a context effect that was due to surrounding information, but the surrounding information in this case was the items being compared themselves! It appears that when objects are compared, each provides a context by which the other is interpreted (also see Hofstadter,

1995). This can result in objects being given radically different, occasionally contradictory, interpretations. Subjects were presented with pairs of objects and listed features that were common or distinctive to the objects. The comparisons were organized around triads of objects (two examples are shown in Figure 2). Half of the subjects saw Object B paired with Object A, and the other half of subjects saw Object B paired with Object C. When Object B was ambiguous, and Objects A and C each instantiated one of B's interpretations, then B's interpretation was shifted *toward* its comparison partner. For example, when subjects compared Object B from the top triad in Figure 2 to Object A, subjects most often mentioned as a shared feature something akin to "both objects have a line crossing two triangles." When prompted for differences, subjects most often responded, "one object has a straight line while the other's line is bent." Both the shared similarities and differences reflect that subjects were interpreting ambiguous Object B so as to be similar to Object A. Object B, when it was compared to Object C, was interpreted as involving three triangles, an interpretation that is unambiguous for Object C. Thus, when

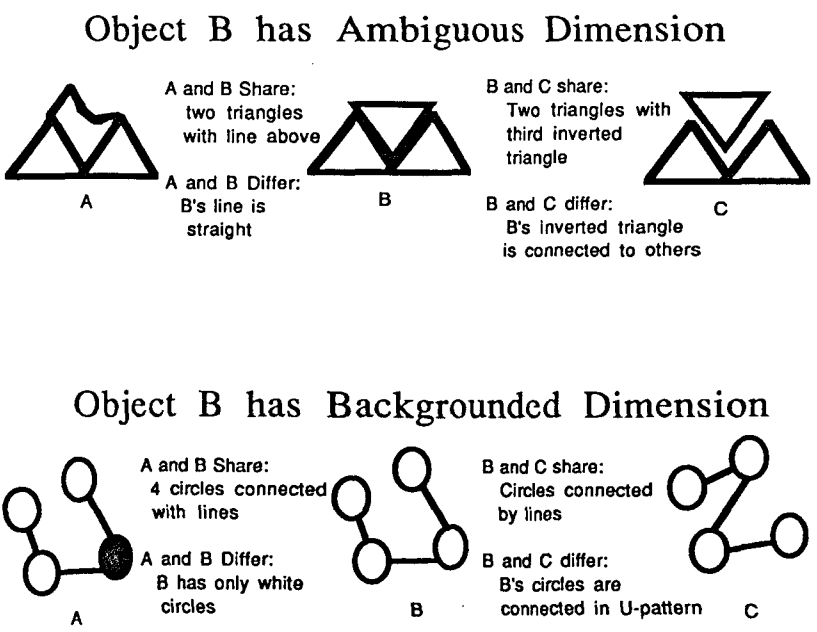

Figure 2. Sample stimuli used by Medin, Goldstone, and Gentner (1993). In the top set, the middle item has two possible interpretations. In the lower set, the middle item has no ambiguity, but has properties that are likely to be backgrounded when it is presented in isolation.

Object B was ambiguous, it was interpreted in a manner that would raise its similarity to its partner in comparison.

A very different pattern of results was found when Object B was not ambiguous, but rather possessed features that were not likely to be considered unless variation along them was present (see also Thibaut, 1991). For example people are not likely to think of Object B as possessing only white circles in the bottom triad of Figure 2. It probably would not even occur to people to think of the color of B's circles, unless contrasting objects were presented that had different colors (Garner, 1962). Consistent with this hypothesis, our subjects did not often list "all white circles" as a common feature between Objects B and C. The whiteness of B's circles was clearly more salient when B was compared to A. In general, the results from Medin et al. indicated that when subjects could interpret one object to be similar to another object along an ambiguous dimension, then subjects would give objects dimension values that heightened their similarity. When this was impossible, then subjects were very likely to represent of the objects in terms of their differing dimension.

Context That is Created by Subjects When There is None

The most radical possible context dependency would occur if, even when the context of the comparison remains constant, comparisons themselves can evoke their own context, in the form of a *contrast set* (Kahneman & Miller, 1986; Lehrer & Kittay, 1992). A contrast set is the set of alternatives brought to mind when a stimulus is presented. Even in situations that involve only two items being compared, different contrast sets may emerge depending on what dimensions are foregrounded or highlighted by the comparison. The creation of different contrast sets for different comparisons can potentially produce violations of the intuitive and pervasive assumption of monotonicity. According to the assumption of monotonicity, adding a common feature to two items should never decrease their similarity, and adding a unique feature to one of the items but not the other should never increase their similarity.

However, in some circumstances, adding unique features may change the salience of a previously backgrounded (ignored because of lack of variation) dimension of the comparison pair, and therefore also the contrast set evoked by the alternatives. In two sets of experiments, we (Goldstone, Medin, & Halberstadt., 1995; Medin et al., 1993) have argued that, in some comparisons, a set of alternatives is spontaneously invoked, implicitly or explicitly, and these alternatives represent the range of features the comparison objects could assume. Consider the top pair of shapes labeled A and B in Figure 3. The contrast set for this comparison would likely be other shapes, perhaps with similar regularity and angularity. The dimension on which the objects differ (i.e., orientation) is salient, while the many ways in which they are similar (e.g.,

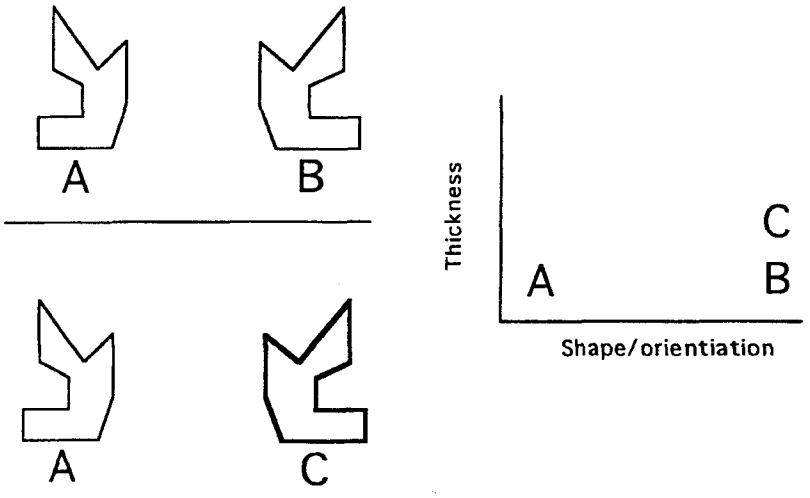

Figure 3. Sample stimuli from Goldstone, Medin, and Halberstadt (1995), demonstrating nonmonotonicities in similarity judgments.

thickness of lines, color, size on the page, texture etc.) are backgrounded — they are not considered relevant for the judgment.

The spontaneous, idiosyncratic evocation of contrast sets can lead to nonmonotonicities in the following way. Consider the second pair of curves, A and C, in Figure 3. A unique feature, line thickness, has been added to C, and the similarity relative to the first pair should therefore decrease according to monotonicity. However, the contrast set for the comparison may have also changed, expanded to include shapes of different thicknesses. In the context of the thick line, it becomes apparent that the shapes could vary on a second dimension, thickness. Yet on this new dimension the two shapes are relatively close; their lines vary only slightly within the range of possible line thicknesses defined by the contrast set. The first pair lies at the extremes of its contrast set because the shapes have opposite orientations, but the second pair are relatively similar in the expanded set. Thus, although a unique feature has been added in the second comparison, it is possible to predict that A and C will receive a higher similarity estimate than will A and B. The abstract characterization of the experimental logic is shown on the right side of Figure 3. Items A and B differ considerably on the horizontal dimension; items B and C have the same value on this horizontal dimension. Items A and B have identical values on the second, vertical dimension. Items A and C have slightly different values on this dimension. If this vertical dimension is a normally backgrounded dimension such as thickness, then subjects who are only given A and B to compare may not even consider their similarity on this dimension when evaluating their

similarity. However, when given the comparison between A and C, a second group of subjects may consider the vertical dimension because of variation along it among the compared items, and increase their similarity estimates accordingly (reasoning that "A and C may be far on the horizontal dimension, but they are quite close on the vertical dimension, so I will give them an intermediate similarity rating").

We employed the design shown in Figure 3 to construct stimuli that we predicted would violate the assumption of monotonicity. Subjects judged the similarity of one pair (A-B or A-C) of each triplet. Results indicated that reliable nonmonotonicities were obtained in some specific conditions (when the Stimulus A was presented first); subjects who were shown the A-C stimuli, gave them a higher similarity rating than subjects who were shown the A-B stimuli. However, when B and C were both presented during the comparison, then strong monotonicities were found (A was more similar to B than C). Simultaneously presenting B and C insures that the same context will be used for the A-B and the A-C comparisons. These results are consistent with theories that suggest that different standards are used for different comparisons (Kahneman & Miller, 1986). Even when context was pared to a minimum by presenting subjects with only single comparisons from unrelated sets, the dimensions that were considered depended on the two compared objects themselves. Context effects seem to be the rule rather than the exception.

The Stability of Similarity

If stimuli can create their own contexts, then it may seem that context effects are utterly unavoidable. Goodman's (1972) position that similarity is too vague and flexible to provide a useful explanation is understandable. If similarity is so flexible that context can cause violations of basic assumptions such as monotonicity and transitivity, then one might conclude that "all bets are off" as far as similarity is concerned. In this section, I hope to dissuade the reader from adopting this perspective. Despite the evidence above, similarity is not infinitely flexible, there are limits to contextual influences, and the flexibility that similarity manifests is also highly systematic.

Pragmatically, similarity is very useful notion. Similarity ratings can predict categorization (as the similarity of A to members of Category X increases relative to other categories, so does the probability of placing A in Category X; Nosofsky, 1986), inferences about an object (as the similarity of A to X increases, so does the likelihood of drawing inferences from properties of X to properties of A; Smith, Shafir, & Osherson, 1993), and memory retrieval (as the similarity of A to X increases, so does the likelihood of a presentation of X causing A to be retrieved from memory; Gentner, Ratterman, & Forbus, 1993;

Ross, 1987). In addition, although different methods of measuring similarity (ratings, same/different judgments, confusions between items, etc.) do not always converge, the correlation is usually fairly high (typically, r=.75 between same/different judgments and ratings). Thus, from the outset, it is apparent that similarity is at least stable enough to be somewhat consistent from one similarity task to another, or from similarity tasks to other cognitive tasks that depend on similarity.

The Primitive Appeal of Overall Similarity

One of the major effects of context is to alter how important a dimension is for a comparison. To take just a few examples, Tversky's diagnosticity effect (1977), the observed nonmonotonicities and intransitivities in similarity, the influences of categorization on perceptual discriminations, and the extension effect (Sjoberg, 1972; Tversky, 1977), can all be described as contextual influences on how much weight particular dimensions receive during a similarity computation. Research on perception suggests that these dimension-weighting accounts may be limited in a fundamental way. A number of researchers have argued that object perception is often "holistic." Object perception is holistic to the extent that it does not rely on analysis of objects into separate aspects, but rather is based on overall similarity. If holistic, overall perception of objects is natural, then there are, by necessity, limits to how influential contextual effects of the dimension-weighting variety can be.

In fact, there is good evidence to suggest that assessments of overall similarity are natural and perhaps even "primitive." Evidence from children's perception of similarity suggests that children are particularly likely to judge similarity on the basis of many integrated properties rather than analysis into dimensions. Even dimensions that are perceptually separable are treated as fused in similarity judgments (Smith & Kemler, 1978). Children under five years of age tend to classify on the basis of overall similarity and not on the basis of a single criterial attribute (Keil, 1989; Smith, 1989). Children often have great difficulty in identifying the dimension along which two objects vary even though they can easily identify that the objects are different in some way (Kemler, 1983). Smith (1989) argued that it is relatively difficult for young children to say whether two objects are identical on a particular property, but relatively easy for them to say whether they are similar across many dimensions.

There is evidence that adults also often have an overall impression of similarity without analysis. Ward (1983) found that adult subjects who tended to group objects quickly also tended to group objects like children, by considering overall similarity across all dimensions instead of maximal similarity on one dimension. Likewise, Smith and Kemler (1984) found that adults who were given a distracting task produced more judgments by overall similarity

than subjects who were not. Several of the cited researchers have argued for a primitive similarity computation that is used when cognitive resources are limited due to age, level of intelligence, distraction, or speed. The essential characteristic of this primitive similarity computation is that it considers a broad number of properties simultaneously. Brooks (1978) argued that judging category membership by overall similarity is an often used strategy, particularly when the category members are rich and multi-dimensional and the category rules are complicated. One may think of holistic processing as just one specific type of strategy that may be invoked when useful, but it is somewhat misleading to consider it a task-tailored *strategy* because it is typically used when little information is available, it is applied across many tasks, and it often is applied even when it is inappropriate.

The above evidence suggests that the most basic similarity computation may not involve analysis of compared entities into parts. Rather, it appears to involve overall similarity across many dimensions. To the extent that similarity is determined by many dimensions, it is less subject to drastic context-driven changes. Some contexts may highlight one dimension while other contexts highlight another, but if basic similarity judgments typically involve all perceptually salient dimensions, then contextual effects will not often drastically alter similarity. Furthermore, to the extent that compared entities are not analyzed into dimensions, it will be difficult to selectively highlight just a single dimension. Even if a context could perfectly bias subjects to attend just a single dimension, if that dimension cannot be successfully separated from the others, or if the segregation process takes substantial time and effort, then the other dimensions will "come along for the ride" when the dimension is attended.

Obligatory Assessments of Similarity

The strong version of the claim for context-driven similarity is that there is no single "generic" similarity between two entities. By this view, the similarity of two entities can be any arbitrary value, depending on one's context and the features one selects (Goodman, 1972; also see parts, but not all, of Murphy & Medin, 1985). Evidence against this strong claim for context-sensitivity comes from psychological evidence suggesting the obligatory use of a relatively stable, context-independent similarity (Goldstone, 1994c). In many situations, even when the correct strategy for solving a task is given to subjects, they still resort to using "generic" similarity. Allen and Brooks (1991) gave their subjects a simple rule for categorizing cartoon animals ("an animal belongs to the category if it has at least two of the following properties: long legs, angular body, and spots"). Despite knowledge of this perfect rule, subjects still categorized new animals on the basis of their similarity to old, previously

shown animals rather than simply using the rule. Subjects seem not have been able to ignore similarities between old and new animals.

This inability to ignore similarities occurs at a less cognitive, more perceptual level as well. When subjects were asked to make same/different judgments on a particular dimension (e.g., color), their response times were influenced by similarity on irrelevant dimensions (Egeth, 1966). The more dissimilar two objects were on irrelevant dimensions, the longer it took subjects to respond "same" on the relevant dimension.

Dimensions that are generally used for similarity comparisons tend to be used even when they are not relevant. Gentner and Toupin (1986) and Ross (1987) found that superficial similarities were used when solving tasks even when these similarities were irrelevant. For example, in solving mathematical word problems ("Five technicians are given 24 computers...."), subjects were influenced by the previous solution of a problem if it involved the same superficial cover story (e.g., both problems involved computers). These similarities were irrelevant, and were known to be irrelevant by the subjects, because the solution to the word problems depended only on the mathematical equations used for their solution, and not their content. Likewise, Sadler and Shoben (1993) showed that subjects' similarity ratings were influenced by features that were relevant for a generic context but not the particular task-defined context. Even when subjects were told to base their similarity on a particular dimension, there was an intrusion of generic dimensions.

The general conclusion to be drawn is that task-defined contexts do not completely determine similarity. Even when a task should be based on only a single property, similarity along other properties leak into people's judgments. The properties that intrude on judgments provide cues as to what is the "generic" similarity of two objects. The rule of thumb is that if the similarity of two objects along property X is influential in a task even when it is irrelevant, then property X is part of the context-free, generic similarity of the two objects (assuming that X is perceptually separable from the task-relevant properties). The processing of irrelevant similarities is obligatory in the sense that consideration of these similarities results in performance that is slower or less accurate than it would be if they were not considered. The similarity of two objects, thus, has more stability than one might imagine from the earlier demonstrations of context-sensitivity. People can adapt to the requirements of a particular task, but they also cannot completely ignore default or generic similarity unless considerable effort is employed.

Steps Toward a Resolution

The end result of reviewing the first two sections of this paper may simply be to convince readers that similarity is somewhat flexible and somewhat constrained. While undoubtedly correct, this conclusion does not take us far in understanding the function of similarity in cognition. By exploring the different functions of similarity, I believe that we can gain a partial understanding of why similarity exhibits its pattern of flexibility and stability.

Default and Directed Similarity

Part of the reconciliation between similarity's stability and flexibility may involve a distinction between default and directed similarity. Default similarity is used when little specific information is known about an entity. As its name implies, unless special-case information is known about the entity, it will be compared to other entities according to a similarity computation that has proven useful earlier. Directed similarity is used when one is in a more informed position. The relatively stable application of a more-or-less generic similarity, then, reflects situations where default similarity is used. The observed contextual dependencies involve directed similarity. Table 1 presents a comparison between default and directed similarity across five dimensions.

Sophistication. Generally speaking, default similarity is less sophisticated than directed similarity. The act of tailoring similarity to a particular purpose is sophisticated in that it requires not taking the standard, automatic route. The use of default, relative to directed, similarity is expected to be greatest when little time is allowed for judgment, when children rather adults are tested, or when a person's attention is distracted by another task. Just as one relies on stereotypical information rather than specific, individual information when one is cognitively pressured (MaCreae, Milne, & Bodenhausen, 1994), so cognitive pressures bias one away from specifically tailored similarity toward more generic similarity.

Information. If one has little information about the entities being compared or the properties that one is interested in, then directed similarity is likely to be used. In Smith et al's (1993) category induction tasks, subjects were given questions such as "Cows have a high concentration of Potassium in their blood. Therefore, mice have a high concentration of Potassium in their blood. How strong is this argument?" In these situations, subjects have little knowledge about the causes of high Potassium concentration in blood. Smith et al's results indicated that subjects resort to using a fairly generic similarity calculation between cows and mice (and information about their common category membership) in order to assess the strength of the argument. Arguing for a more directed similarity, Heit and Rubinstein (1994) showed that anatomical properties

Table 1
Characteristics of Default and Directed Similarity

Comparison dimension	Default similarity	Directed similarity
Sophistication	Simple	Sophisticated
Information	Little information	Substantial information
Categories	Taxonomic	Ad-hoc, metaphors
Course of tailoring	Long term	Short term
Inferential use	General	Specific

were more likely to be inferred when animals shared anatomical, rather than behavioral, similarities, and *vice versa*. Hypothetically, if a specific, familiar property was probed, as with "Wild boars are dangerous. Are pigs?", then very specific similarity of the animals along this property would be used to make the induction rather than their generic similarity; although the two animals are generically similar, this similarity would not be used because specific information about the probed property is available. Thus, specialized, directed similarity is used to the extent that people have specific information about objects and their properties.

Categories. Default, stable similarity provides us with our default, stable categories, whereas directed similarity is used for dynamically created, creative categories. Taxonomic categories include *table*, *dog*, and *car*. Objects that belong in taxonomic categories have many properties in common; their overall similarity across many properties is high (Rosch, 1975; Rosch & Mervis, 1975). The taxonomic category membership of an object seems to be automatically determined (Barsalou & Ross, 1986). On the other hand, things that belong to the same ad hoc category (*things to take out of a burning house*, Barsalou, 1983) or abstract analogies or metaphors (e.g., events *in which a kind action is repaid with cruelty, metaphorical prisons*, and *problems that are solved by breaking a large force into units that converge on a target*) are not similar overall. An unrewarding job and a relationship that cannot be ended may both be metaphorical prisons, but this categorization is not established by overall similarity. The situations may seem similar in that both conjure up a feeling of being trapped, but this feature is highly specific, and similarity must be *directed* toward this property if the situations are to be deemed similar. The sensitivity of explicit similarity ratings to analogical similarity (Gentner, 1989; Gentner & Markman, 1995; Goldstone, 1994b; Holyoak & Thagard, 1989; Markman & Gentner, 1993; Medin et al., 1993) and ad-hoc category membership (Barsalou, 1982) attests to the degree to which similarity can be directed toward a particular type of comparison or property.

Inferential use. As suggested by the above differences between default and directed similarity, the two similarities both provide useful roles in inferences, but these roles are different. Default similarity allows general inferences to be drawn. If we see one crocodile eat a chicken, default similarity allows us to infer that other crocodiles will probably eat chickens, and probably alligators will as well. We may not selectively weight particular properties of the crocodiles and alligators in making our inference for three reasons. First, we may not know how to break the objects into parts or properties. Materials from psychological experiments are often easy to analyze in terms of parts, but real world objects usually do not wear their segmentations on their sleeves (Schyns, Goldstone, & Thibaut, 1995). Second, even if one can break the objects into parts, one may not know which similarities are relevant for an inference and which are irrelevant. If this discrimination cannot be made, attending to all parts is the most reasonable course. Third, there may be no need to analyze the objects into parts and selectively weight the parts on the basis of their relevance. When objects are highly similar on virtually *all* of their properties (e.g., one crocodile compared to another), then selective weighting will have virtually no effect on the resulting inferences that are drawn. In such a case, all parts point to the same conclusion (Brooks, 1978). For these three reasons, it is not always necessary or possible to tune property weightings to particular tasks. The end result is that default weightings of parts, rather than context-tuned weightings, are used for guiding inferences from one entity to another. Although general inferences are generally useful, specific inferences must also be drawn. For these, directed similarity is needed. When a child learns that a hook permits them to retrieve candy, then they may be able to infer that another device will also permit this. However, not all hook-to-device similarities are relevant, and even children seem to understand this (Brown, 1990). In nonmonotonic reasoning systems that have default assumptions (e.g., birds fly), specific information can be used to overrule these defaults (e.g., penguins are birds, but penguins do not fly). Likewise, directed similarity is used in situations where default similarity is overruled by background knowledge or idiosyncratic intentions.

Course of tailoring. The arguments for a default, generic similarity may be misinterpreted as arguments for strong realism and an objective, observer-independent metric of similarity. Default similarity as conceptualized here does not require realism of the type that would hypothesize objective similarities between objects. Default and directed similarities are both tailored or custom-tuned to the observer. However, the tuning of directed similarities is short term and temporary. Directed similarity can highlight different aspects of a comparison in the course of an hour-long experiment. When asked to compare *skin* to *bark*, their metaphorical similarity (both are used to cover and protect an organism) is highlighted. Ten minutes later, if *skin* is compared to *hair*, then the more concrete "possessed by human bodies" aspect of *skin* is highlighted (Medin et

al., 1993). The retuning of property importance is quickly executed, but not particularly long lasting.

Conversely, impressions of default similarity are tuned to the judging organism, but much of this tuning is conducted on an *evolutionary* time scale. Mother nature rewards organisms that perceive similarities between entities that have similar survival value for the organism. She lets such organisms live to propagate. The default similarities that we are born perceiving are still tuned to the organism's life tasks, broadly construed. Other default similarities are tuned over the life course of the organism. Speakers of French, because they learn that /r/ and /l/ do not have the same uses in their language, learn to perceive dissimilarities between these phonemes that Japanese speakers do not. Although these similarities are tuned, they become part of the default similarity that is used by adult speakers — it takes effort for speakers *not* to take them into account.

Given that both default and directed similarities are tailored, the distinction is properly viewed as a continuum rather than a dichotomy. A directed similarity that has been noticed repeatedly by an individual over a lifetime may become part of the default repertoire of the organism. When an individual has a great deal of time, information, and resources to dedicate, directed similarities will tend to be used. As these resources are gradually reduced, the comparison will revert progressively to a more default similarity.

Perceiving Conceptual Similarity

The proposed distinction between default and directed similarities may remind readers of Quine's (1977) distinction between innate similarity and scientific categories. For Quine, people and cultures begin existence with an innate sense of similarity, but with increasing sophistication, they discard this innate similarity in favor of similarity that has a scientific basis: "...it is the mark of maturity of a branch of science that the notion of similarity or kind finally dissolves" (p. 35). For example, whales and fish may originally seem more similar than whales and horses, but our scientific knowledge can direct us to physiological similarities that reverse this judgment.

In the current view, default similarity is never supplanted by "smarter" similarities, and we would not want it to be. Default similarity allows inferences to be drawn that are beyond the scope of directed similarity. Default similarity does not mislead us; it is explicitly *designed* to lead us to see relations between things that often function similarly in our world (Medin & Ortony, 1989).

One reason why default similarity is not a source of information to be discounted or "risen above," as is Quine's innate similarity, is that default similarity, in addition to gradually changing with evolutionary pressures and an individual's life course, provides the grounding for more sophisticated

similarities. Directed similarities leap off from the starting point established by default similarity. Default similarity is the central camp from which scouts are sent. Once default similarity allows us to see the similarities between crabs and lobsters, directed similarity can make more remote connections, linking crabs with spiders under the category *arthropods*. A biologist would be much more likely to notice genetic similarities between dogs and wolves than between dogs and roses. People, with good reason, expect their default similarity assessments to provide good clues about where to uncover directed, nonapparent similarities (Goldstone, 1994c; Medin & Ortony, 1989). As default similarity grounds directed similarities, so there is an influence in the opposite direction as well. Directed similarities, if they regularly follow a systematic pattern, become default similarities over time. Similarities that were once effortfully directed, become *second nature* to the organism. Roughly speaking, this is the process of *perceiving* what was once *conceptual* similarity. At first, the novice mycologist explicitly uses rules for perceiving the dissimilarity between the pleasing Agaricus Bisporus mushroom and the deadly Amanita Phalloides. With time, this dissimilarity ceases to be effortful and rule-based, and becomes perceptual and phenomenologically direct. When this occurs, the similarity becomes default, and can be used as the ground for new directed similarities. In this way, our cognitive abilities gradually attain sophistication, by treating territory as ground that once made for difficult mental climbing. To return to the metaphor used earlier, the "scouts" of directed similarity may always have default similarity as their point of departure, but if one of the scouts finds a particularly promising site, the entire camp may move. In this manner, default similarity provides a ground for more sophisticated similarities, but is, in turn, influenced by the sophisticated similarities that are noticed. Given these comments, a very directed similarity can be suggested: "Directed similarity is to default similarity as the artistic avant-garde (literally: 'front line') is to mainstream culture." Both domains have a conservative, slow-moving force, and an exploratory vehicle of change. The avenues explored by the avant-garde and by directed similarity are usually never followed by their more conservative partners. In music, Alois Haba's microtonal music of the twenties has never procured mainstream favor. In cinema, the split-screen technique of multiple simultaneous action, used by Abel Gance's 1927 "Napolean," has inspired few directors (and mostly other avant-garde directors at that, such as Andy Warhol).

A small number of the explored avenues are adopted by the mainstream. In music, Igor Stravinsky's dissonances, originally considered utter noise by many, are part of our mass musical consciousness in 1995. In cinema, the use of fragmented quick-takes was hopelessly perplexing to most audiences of 1920. A casual glance at contemporary music videos testifies to the entrenchment of these techniques in the contemporary mainstream.

Likewise, most directed similarities are created within a particular, short

term context, but a minority will "catch on" sufficiently to change the nature of default similarity. There is a continuum between default and directed similarities — the continuum of context-independence. As the similarity between two entities is noticed in increasingly many contexts, it moves from the directed to the default pole. It is important that our similarity assessments partake from both poles. Default similarity allows us to draw inferences from one object to another that are likely to be correct across many contexts. Directed similarity allows us to adjust our judgments to short term contextual demands. As with artistic culture, tradition provides grounding to otherwise capricious change, and change provides adaptability to otherwise stagnant tradition.

References

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General, 120*, 3-19.

Asch, S. E. (1952). *Social psychology*. New York: Prentice-Hall.

Attneave, F. (1950). Dimensions of similarity. *American Journal of Psychology, 63*, 516-556.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition, 10*, 82-93.

Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition, 11*, 211-227.

Barsalou, L. W., & Ross, B. H. (1986). The roles of automatic and strategic processing in sensitivity to superordinate and property frequency. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 1*, 116-134.

Brooks, L. R. (1978). Non-analytic concept formation and memory for instances. 169-211. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*, Hillsdale, NJ: Erlbaum.

Brown, A .L. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science, 14*, 107-133.

Burns, B., & Shepp, B. E. (1988). Dimensional interactions and the structure of psychological space: The representation of hue, saturation, and brightness. *Perception and Psychophysics, 43*, 494-507.

Diamond, J. (1966). Classification systems of primitive people. *Science, 151*, 1102-1104.

Egeth, H. E. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics, 1*, 245-252.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity, analogy, and thought*. Cambridge, MA: Cambridge University Press.

Gentner, D., & Markman, A. B. (1995). Similarity is like analogy. In C. Cacciari (Ed.), *Similarity in language, thought, and perception* (pp. 111-148). Brussels: Brepols.

Gentner, D., Ratterman, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer:

Separating retrievability from inferential soundness. *Cognitive Psychology, 25,* 524-575.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10,* 277-300.

Goldstone, R. L. (1994a). influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General, 123,* 178-200.

Goldstone, R. L. (1994b). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 3-28.

Goldstone, R. L. (1994c). The role of similarity in categorization: Providing a groundwork. *Cognition, 52,* 125-157.

Goldstone, R. L., Medin, D. L., & Halberstadt, J. (1995). *Similarity in context.* Indiana University Cognitive Science Technical Report.

Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (Ed.), *Problems and projects.* New York: The Bobbs-Merrill Co.

Gregson, R. A. M. (1975). *Psychometrics of similarity.* New York: Academic Press.

Harnad, S. (1987). *Categorical perception.* Cambridge: Cambridge University Press.

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

Hofstadter, D. (1995). *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought.* New York: Basic Books.

James, W. (1890/1950). *The principles of psychology* (Vol. 1). New York: Dover.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93,* 136-153.

Kelly, M. H., & Keil, F. C. (1987). Metaphor comprehension and knowledge of semantic domains. *Metaphor and Symbolic Activity, 2,* 33-51.

Kemler, D. G. (1983). Holistic and analytic modes in perceptual and cognitive development. In T. J. Tighe & B. E. Shepp (Eds.), *Perception, cognition, and development: Interactional analyses* (pp. 77-101). Hillsdale, NJ: Erlbaum.

Lehrer, A., & Kittay, E. F. (1992). *Frames, fields and contrasts.* Hillsdale, NJ: Erlbaum.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25,* 431-467.

MaCreae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66,* 37-47.

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, MA: Cambridge University Press.

Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology, 20,* 158-190.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289-316.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115,* 39-57.

Quine, W. V. O. (1977). Natural kinds. In S. P. Schwartz (Ed.), *Naming, necessity, and natural kinds.* Ithaca, NY: Cornell University Press.

Rosch, (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 303-322.

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

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 629-639.

Sadler, D. D., & Shoben, E. J. (1993). Context effects on semantic domains as seen in analogy solution. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 128-147.

Schyns, P. G., Goldstone, R. L., & Thibaut, J-P (1995). *The development of features in object concepts.* Indiana University Cognitive Science Technical Report #133. Bloomington, Indiana.

Shepard, R. N. (1962) The analysis of proximities: Multidimensional scaling with an unknown distance function. Part 1. *Psychometrika, 27,* 125-140.

Sjoberg, L. (1972). A cognitive theory of similarity. *Goteborg Psychological Reports, 2*(10).

Smith, E. E., Shafir, E., & Osherson, D. (1993). Similarity, plausibility, and judgments of probability. *Cognition, 49,* 67-96.

Smith, L. B. (1989). From global similarity to kinds of similarity: The construction of dimensions in development. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 146-178). Cambridge, MA: Cambridge University Press.

Smith, L. B., & Kemler, D. G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology, 10,* 502-532.

Smith, J. D., & Kemler Nelson, D. G. (1984). Overall similarity in adults' classification: The child in all of us. *Journal of Experimental Psychology: General, 113,* 137-159.

Thibaut, J. P. (1991). *Recurrence et variations des attributs dans la formation de concepts.* Unpublished doctoral dissertation, University of Liege, Liege.

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

Ward, T. B. (1983). Response tempo and separable-integral responding: Evidence for an integral-to-separable processing sequence in visual perception. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 103-112.

Whorf, B. L. (1941). Languages and logic. In J. B. Carroll (Ed.), *Language, thought, and reality: Selected papers of Benjamin Lee Whorf* (pp. 233-245). Cambridge, MA: MIT Press (1956).

Received February, 1995