Weighting in Similarity Judgements: Investigating the "MAX Hypothesis"

Ulrike Hahn and **Nick Chater** and **Rachel Henley**
Department of Experimental Psychology
University of Oxford
South Parks Rd., Oxford OX1 3UD, England
{ulrike,nick}@psy.ox.ac.uk

Abstract

Most models of similarity assume differential weights for the represented properties. However, comparatively little work has addressed the issue of how the cognitive system assigns these weights. Of particular interest to the modelling of similarity are factors which arise from the comparison process itself. One such factor is defined by Goldstone, Medin & Gentner's (1991) 'MAX Hypothesis'. We present a series of experiments which clarify the main components of 'MAX' and examine its scope.

Introduction

Similarity as an explanatory principle is so ubiquitous in cognition, that stressing its importance borders on banal. Similarity is central to theories of categorization, learning, memory, and problem-solving as virtually any paper on the topic documents (see Goldstone (1994a) for an overview). Its study has two main goals. First is the discovery of the function according to which the various matches and mismatches that the two objects under consideration exhibit are combined into a single similarity judgement. This has been the main focus of similarity research in the past, exemplified by spatial models of similarity or Tversky's contrast model (Shepard, 1962; Tversky, 1977). The second goal, is discovering how the *relevant* properties[1] are determined in the first place: According to what factors are the relevant properties selected from the infinite set of properties any object possesses, and how are they assigned the differential weights most models assume? Given the profound representation-dependence of similarity (Goldstone, Medin & Gentner, 1991; Hahn & Chater, 1996), this latter goal is at least as important as discovering of cognitively plausible similarity functions, but it has remained largely outside the scope of current experimentation and modelling. For theorists, the lack of constraints on which properties are represented gives 'similarity' a flexibility that makes the notion almost vacuous. This has long been at the heart of criticisms of similarity-based explanations in cognition (Goodman, 1972; Medin & Wattenmaker, 1987; Goldstone, Medin & Gentner, 1991).

Though the centrality of selection and weighting to theories of similarity is widely acknowledged, we yet know little about it. In general, two types of influences can be distinguished: *knowledge-based* and formal, non-knowledge-based factors which we will refer to as *process principles* (see also Goldstone, Medin & Gentner (1991); Hahn & Chater (1996)). As an example of knowledge-based factors, one can think of the impact that 'theories' (scientific or informal) have on what we consider to be important properties of an object (Medin & Wattenmaker, 1987). Process principles, on the other hand, are influences arising from the comparison process itself. A prominent example is Tversky's 'diagnosticity principle' (Tversky, 1977). In assessing the similarity structure of a set of objects we increase the weight of properties which enable further subdivisions of the set, and depress those which are common to all members of the set in question. Hence, for example, the feature 'real' has little diagnostic value in the set of 'mammals' since it is common to all. When the set is extended to include Pegasus, unicorns, and mermaids, however, the diagnostic value, and thus weight, of 'real' is considerably increased, thus indicating a systematic relationship between feature weights and sub-clusters of objects within a set.

Other examples of process principles proposed and experimentally investigated in the psychology literature are the 'focussing hypothesis' (Tversky, 1977) and the role of 'structural alignment' (Goldstone, 1994b). Process principles must be considered central to the study of similarity, not only because they embody constraints on the all-important feature selection and weighting process, but because their general, formal nature makes them candidates for inclusion in future, more explanatory, models of similarity.

It is such a process principle—Goldstone, Medin & Gentner's "MAX-hypothesis" (1991)— that we wish to discuss in this paper. In the following, we will review the MAX hypothesis, present a more explicit version of its various components and present new experimental results. Finally, we discuss implications of these results for the modelling of similarity.

The "MAX-hypothesis"

Goldstone et al. (1991)

The MAX-hypothesis was first put forth and experimentally investigated by Goldstone, Medin & Gentner (1991). It proposes that property weights are, in part, dependent on the particular *type* they belong to as well as the extent to which this type dominates in the stimuli of the similarity comparison.

[1] We use 'property' in the widest possible sense to encompass attributes or relations, binary features or continuous dimensions.

The two types in question are *attributes*, i.e. one-place predicates (e.g. 'color(X)'), and *relations*, that is, predicates with an arity of two or more (e.g. 'larger-than(X,Y)'). Specifically, Goldstone et al. claim

> "(1) attributional similarities are pooled together,and relational similarities are pooled together, and (2) the weight that a similarity has on the final similarity judgement increases with the size of the pool to which it belongs" (p.228.)
>
> *or:*
>
> "MAX claims that relations and attributes are psychologically distinct, that similarities are classified as relational or attributional types, and that people attend more to similarities belonging to whichever similarity type is greatest" (p.228.)

This can be illustrated with a brief example. Imagine a group of five stimuli *A,B,C,D,* and *T* as illustrated in Figure 1. Subjects are asked to assess the respective similarities of *A,B,C,* and *D* to the "target stimulus" either through direct judgements (e.g. 'how similar are *A* and *T*?') or a series of forced choices (e.g. 'which is more similar to *T*, *A* or *B*?').[2]

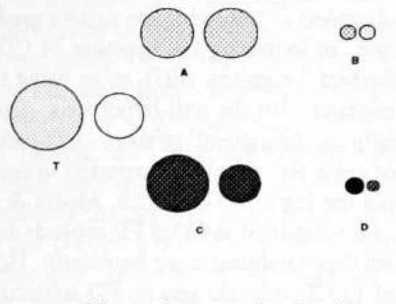

Figure 1: example stimuli

One strategy subjects could adopt is to assign differential weights to the component properties (relations or attributes) *uniformly* for the entire stimulus set; e.g. a match in "size' is twice important as the relation "bigger than". By contrast, the MAX-hypothesis predicts that the weight of relations is differentially boosted in stimuli in which relational matches to the target dominate (here *D*) and are depressed where attribute matches dominate (here *A*). In other words, 'bigger than' is more important in *D* which also shares the target's 'darker than' relation, than it is in *A*. This means that weighting of a property depends not only on its intrinsic salience, but is determined in systematic interaction with the other properties of the stimulus.

The MAX-hypothesis was supported in a series of experiments by Goldstone, Medin & Gentner (1991). Our interest has been to tease apart various claims implicit in the MAX-hypothesis as stated there, and, on the basis of this, to investigate its generality. We begin with our theoretical considerations.

Spelling out the underlying claims

The MAX-hypothesis involves two claims: first, that relations and attributes are 'pooled' and second, that the largest pool is boosted. Both aspects can be regarded independently. The fact that the largest pool is boosted may be due to a more general principle, which we call the "Closest Dimension Principle" (CD), which can apply to single properties and pools alike. Similarly, pooling may not be restricted to relations and attributes.[3] We address these in turn.

1. The Closest-Dimension Principle.
Part (2) of the MAX-hypothesis as quoted above states that the pool which 'maximizes overall similarity' is selectively boosted in weight. We suspected this was an instance of a more general principle —to be called the "Closest Dimension Principle"—which *generally* states that the dimensions exhibiting the greatest similarity to those of the target are boosted. For example, imagine a set of stimuli differing from the target on two dimensions, size and shade. The Closest Dimension Principle predicts that the importance given to 'shade' as opposed to 'size' will be maximal in the stimulus which is *closest* to the target in shade, whereas 'size' will be maximally weighted in the stimulus closest in size.

Boosting the 'largest pool', then, is merely an instance of this principle: the dimensional value of interest is the size of the particular pool. Through pooling, the pooled properties become a single unit to which the CD-principle is applied.

Experiments 1 and 2, described below, investigated the CD principle in non-pooling (i.e. single dimension) situations, both for (singularly treated) attributes and relations.

2. Pooling.
We suggest that the cognitive system may not adopt a principle which applies uniquely to relations and attributes, but may use a more flexible approach. Perhaps the scope of 'pooling' is both wider and narrower than indicated by the MAX-hypothesis. We expect *pools of other types*, that is pooling along lines other than relation/attribute. We also expect cases were pooling according to relations or attributes fails, because these pools are not suggestive in the particular materials. Pooling, we think, may profitably be viewed as a kind of conceptual 'chunking'.

This suggests a wide field of research into the types of pools subjects form, into pool-size, and into the conditions under which pooling occurs. Experiments 3a, b, and c investigate alternative types of pools, whether relations and attributes are necessarily pooled, and whether alternative pools might be preferred over pooling according relations and attributes.

[2] Both forced choice and direct judgements give the same results (Goldstone, Medin & Gentner, 1991).

[3] This possibility is already suggested in Goldstone, Medin & Gentner (1991).

Experimental Results

The CD principle

Principles of Stimulus Design for Experiments 1 and 2 The basic structure of our materials is the same for Experiments 1 and 2 and follows Goldstone, Medin & Gentner's (1991) original experiments. The stimuli come in groups of 5 (as in Fig. 1), one of which is the "target" stimulus which serves as the reference point; subjects are asked to indicate 'which is more similar to *T,A* or *B*?'. For reasons of experimental control, total or overall similarity to the target must be equal for each of the four non-target stimuli, or more precisely *would be equal if the CD principle did not apply*. To this end stimuli are constructed according to the following pattern. For each group there are two properties which are varied, such as 'height' and 'degree of tiltedness'. These properties are manifest in five different levels, distributed as follows: The target stimulus has the maximal level for both properties i.e. a 'dimensional value' of 5/5. Next, Stimulus *A*, has the closest dimensional value for one property and the furthest on the other, i.e. 4/1. Stimuli *B* and *C* have the middle values 3/2 and 2/3; *D*, finally, has 1/4. Hence for all stimuli, the overall "distance" to the target is equal (i.e. five).

The CD principle predicts that 'dimension 1' is selectively boosted in stimulus A and 'dimension 2' selectively boosted in D. This is evidenced as a change in preference of dimension as a criterion for decision. In comparing *A* and *B* to *T*, dimension 1 is viewed as most salient whereas in the judgement *C,D* dimension 2 becomes more salient. Alternatively, subjects could boost the furthest-dimension or merely stick to one dimensional weighting throughout.

Experiment 1: The CD-Principle and Relations
Materials In Experiment 1, the properties of interest were two single *relations* present in the the stimuli. These relations were 'dimensionalized' to fit the above schema. This is clearest with an example: dimension 1, for instance, could be based on the relation 'distance(x,y)' between two component parts of the stimuli. The distance between x and y in the target stimulus forms the base-line, maximal value. We then construct four stimuli with increasing distance between x and y; the distance closest to that found in the target (i.e. the shortest) forms 'level 1' of the schema above, the next closest 'level 2' and so on. The relations manipulated in this way were spacing, displacement, alignment, overlapping of components, occlusion, symmetry, and relative position. A sample group in which relative distance and occlusion were used is shown in Figure 2.[4]

We used 8 different stimulus sets which involved different kinds of component objects (lines, circles, squares etc.) and different pairs of relations as the two manipulated dimensions. The 32 judgements of interest (4 per group) were presented in booklets, with three trials to the page. Each trial consisted of a target and two comparison stiumuli. The subject was asked to indicate which of the two comparision stimuli was most similar to the target by circling it. The order of the comparisons was randomized with the constraint that no two judgements from the same group appear on the same page. The internal order of each comparison pair was randomized and counterblanced between two groups of subjects.

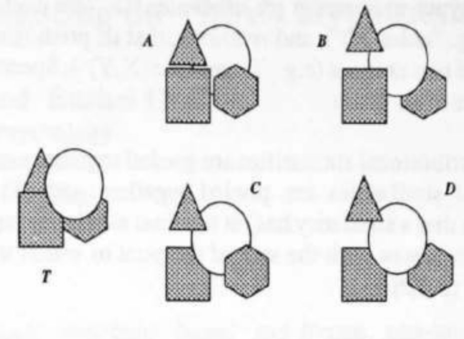

Figure 2: A stimulus group from Exp. 1.

Participants The 14 participants, of which equal numbers were male and female, were members of the departmental subject panel, or students of the university. They represented a wide range of backgrounds and were aged between 18 and 50.

Results The raw data consist of the preferences indicated on the comparisons *A,B; A,C; C,D* and *B,D* for each group. These can be classified as following the pattern predicted by the CD principle, as following the opposite of CD, that is boosting the *furthest* dimension (FD), or as using the same weighting throughout. On the null-hypothesis –assumption of a 'same-weighting-throughout' strategy– CD-patterns and FD-patterns are noise and should be expected in equal numbers. Following the logic of Goldstone, Medin & Gentner (1991), subjects are classified as CD or FD subjects depending on which pattern they produced more frequently. This classification yielded 13 CD-subjects and no FD-subjects. 1 subject had equal numbers of CD and FD patterns and so was left out of the analysis. The hypothesis that there would be more CD subjects than FD subjects was tested using a one-tailed binomial test. The null-hypothesis was rejected at level $p < 0.0005$.

Experiment 2: Attributes Experiment 2 differs from Experiment 1 only in that the two manipulated dimensions were attributes, not relations. The main attributes used were shading, size, and orientation. A sample group, manipulating curvature and line-thickness is shown in Figure 3.

The 18 participants were from the departmental subject panel or postgraduate members of the university. Of these, 12 were CD-subjects, and 3 were FD-subjects. 3 subjects had equal numbers of CD and FD patterns and, thus, dropped out of the analysis. There were significantly more CD-subjects than FD-subjects with the null hypothesis rejected at $p < 0.018$, again using a one-tailed binomial test.

In summary, the data in both experiment seem to confirm the Closest-Dimension principle as an important factor

[4]Letters, here and in the following, are added to aid the reader; they were not part of the actual stimuli.

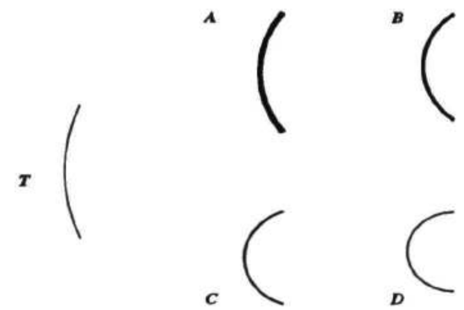

Figure 3: A stimulus group from Exp. 2.

in property weighting.

"Pooling"

Design The point of the remaining three experiments was to investigate the generality of the MAX-hypothesis in two directions: whether other types of properties 'pool' and whether attributes and relations necessarily pool. To address the first issue, we designed objects with properties that could be pooled on the basis of *'theme'*. Themes are groupings such as "colour-related" or "to-do-with-size", which might unite an attribute and a relation in a common pool. For example, the attribute '(actual) size' and the relation 'larger-than' might form the basis of a 'size-related' pool. 'Themes' always involved one attribute and one relation. There were two 'themes' per group, i.e. 2 relations and 2 attributes. Hence, it could be assessed directly whether subjects preferred pooling by theme or by relation/attribute. This can be illustrated with Fig. 1 above. Subjects could pool according to attributes and relations, leading to an increased similarity of stimulus *A* (which matches the target closely on the attributes 'size' and 'shade') and *D* (which matches on the relations 'larger-than' and 'darker-than' between the left and right component) to the target. Alternatively, subjects could form the theme pools 'size-related' and 'colour-related', boosting the similarity of stimulus *B* (matching the target closely in actual shade and the 'darker-than' relation) and *C* (matching closely on actual size and the 'bigger-than' relation). In addition each group had two 'control' stimuli, *E* and *F*, with no (apparent) opportunities for pooling, to allow baseline assessment of whether pooling was occurring at all, and, hence, responsible for any difference in subject's preference between 'theme-pool' objects and 'relation/attribute-pool' objects.

The various properties were distributed across objects according to the scheme in Table 1.

As indicated in the table, the target stimulus for each group has the maximum value for all properties. The overall similarity of the remaining stimuli to the target, disregarding pooling, is held constant. A difficulty is posed by the stimuli which require attribute matches in the absence of relational matches (in particular, object *A*), since two objects cannot completely match attributionally without matching relationally as well. Our strategy was to take middle values or averages in order to approximate this attributional match. The imperfection of this match was balanced by adding a corresponding decrement to the attributional match in the remaining stimuli even where an exact match would have been possible. A sample group, where this can be seen, is given in Figure 4.

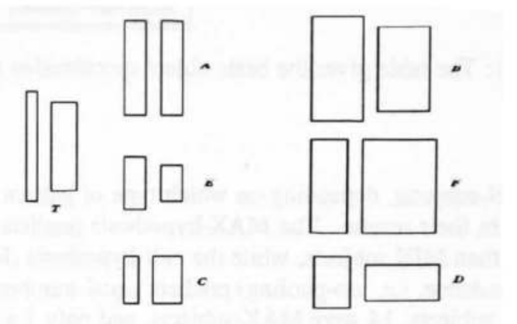

Figure 4: A stimulus group from Exp. 3.

In Figure 4, *A* matches closely in actual size even though it lacks the larger-than relation; [5] The attributional 'match' in actual-size of the theme-stimulus *C* is designed to deviate from the target's values by a corresponding degree, by making both of the component objects slightly larger than in the target.

For maximal experimental control, each subject was asked to give all 12 judgements per group (at 8 groups a total of 96 judgements). In the following, however, we will analyze and present these results as three experiments. All other factors were kept as in Experiment 1 and 2. Participants were 18 members of the university.

Experiment 3a: Pooling According to Themes The first experiment, looks only at the similarity judgements for the 'theme' objects and the two controls, aiming to establish whether pooling according to themes can be found. To this end, the judgements pertaining to the 'theme' stimuli (*B* and *C*) and the two controls are extracted from the 96 judgements. This yields a set which is isomorphic to experiments 1 and 2[6] and which can be analyzed in exactly the same way. What was the single closest dimension in these experiments is now a potential pool. Hence, the MAX-hypothesis applies directly: subjects form pools of properties (here, according to theme) and differentially boost the weight of the properties in the largest pool (hence applying the CD-principle no longer to individual properties but to pools). Following Goldstone, Medin & Gentner (1991), judgements displaying this pattern are classified as MAX. Patterns were the smallest pool is preferred are classified as MIN. Patterns explicable with a single set of weights throughout are classified as "independent". Then subjects are classified as MAX-subjects

[5] Similarly, *A* in Figure 1 closely matches the target in actual colour and actual size, although it lacks the relations darker-than and larger-than.

[6] and now directly corresponds to Goldstone, Medin & Gentner's (1991) experiments except that potential pools are themes instead of relation and attributes.

Possible Combinations of Properties							
	A	B	C	D	E	F	T
relation-*theme1*	-	+	-	+	+	-	+
relation-*theme2*	-	-	+	+	-	+	+
attribute-*theme1*	+	+	-	-	-	+	+
attribute-*theme2*	+	-	+	-	+	-	+

Table 1: The table gives the basic object specification for experiments 3a,3b,3c which allow both relation/attribute and theme pools.

or MIN-subjects, depending on which type of pattern dominates in their results. The MAX-hypothesis predicts more MAX than MIN subjects, while the null-hypothesis (feature independence, i.e. no-pooling) predicts equal numbers. Of the 18 subjects, 14 were MAX-subjects, and only 1 a MIN-subject. 3 subjects had equal numbers of MAX- and MIN-patterns and, hence, were excluded from the analysis. The difference between numbers of MAX- and numbers of MIN-subjects was significant; using a one-tailed binomial test, the null-hypothesis was rejected at $p < 0.0005$.

Experiment 3b: Pooling According to Relations and Attributes The second, Experiment 3b, looks only at the 'relations' and 'attribute' stimuli (*A* and *D* in the table above) and the two control objects, to establish whether Goldstone, Medin & Gentner's (1991) pooling of relations and attributes is replicated with our stimuli. The analysis follows directly that of Experiment 3b. This replication failed; in fact, MIN-marginally exceeded MAX subjects. In actual numbers, we found 7 MAX-subjects, 9 MIN-subjects, and 2 subjects with equal numbers of MAX- and MIN- patterns.

Experiment 3c: Pooling-Preferences - Themes vs. Relations/Attributes In the last of the three, Experiment 3c, we directly compared the two theme-stimuli with the 'relations' and the 'attributes' stimulus. Unsurprisingly, given the previous two results, subjects overwhelmingly preferred pooling according to themes (MAX-subjects: 15, MIN-subjects: 2, equal numbers: 1), with the null-hypothesis (independence, i.e. no pooling) rejected by two-tailed binomial test at $p < 0.002$.

Discussion

We identified two seperate factors implicit in the original formulation of the MAX-hypothesis, the Closest Dimension Principle and the existence of pooling. Our results confirm the CD-principle both for single properties (attributes or relations) and for pools. As regards pooling, our results indicate both that pooling can occur along other lines than relations and attributes and that pooling by relations and attributes can fail. This suggests that, contrary to Goldstone, Medin & Gentner's (1991) assumptions, the relations/attributes distinction is not *per se* cognitively salient enough to evoke pooling. Thus, the criteria determining whether or not a particular pool is salient or suggestive remain to be determined. We are threatened with a certain explanatory regress; the task of determing property weights, i.e. salience, requires determining the salience of potential pools. Only future work can tell whether general principles can be uncovered here.

In the meantime, these results leave the modeller in an awkward position. The Closest Dimension Principle, whether it operates on single dimensions or on pools, has damning consequences for traditional models of similarity such as the contrast model and spatial models of similarity (see Goldstone, Medin & Gentner (1991)). These models omit the feature weighting process, assuming a set of weighted features or dimensions as given. The operation of the CD-principle, however, is *local to the individual comparison*, e.g. target/object1.[7] This means that each individual pairwise comparison requires a different set of weights. There is no single set of weights, even for a group of 5 stimuli as simple as that in Fig. 1. Hence, the contrast model or spatial models of similarity —or any other model which assumes weighted properties as primitives–cannot fit this group of stimuli, but only individual pairwise comparisons. Hence, these models cannot be used to *explain* the similarity structure of even this simple group. They can be extended to include the systematicity in the group only by inclusion of the weighting process itself.

In this respect, however, our results are problematic. They indicate that the conditions under which pooling occurs, may be far more diverse than Goldstone, Medin & Gentner (1991) suggest. Whether they are at all predictable enough to allow inclusion in a formal model remains to be seen.

Acknowledgements

Thanks go to Todd Bailey and three anonymous reviewers who provided helpful comments on earlier versions of this manuscript; remaining errors are, of course, our responsibility. Ulrike Hahn is supported by ESRC grant No. R004293341442.

References

Goldstone, R. (1994a). The Role of Similarity in Categorization: Providing a Groundwork. *Cognition*, 52, 125—157.

[7] please note again that the forced choice design of our experiments is extrinsic to the CD-principle as demonstrated by Goldstone, Medin & Gentner (1991).

Goldstone, R. (1994b). Similarity, Interactive Activation, and Mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 3—28.

Goldstone, R & Medin, D. L. & Gentner, D. (1991) Relational Similarity and the Nonindependance of Features in Similarity Judgements. *Cognitive Psychology*, 23, 222—262.

Goodman, N. (1972). Chapter: Seven Strictures on Similarity. *Problems and Projects.* Indianapolis: Bobbs-Merill Comp.

Hahn, U. and Chater, N. (1996). Concepts and Similarity. *In:* Lamberts, K. & Shanks, D. R. (Eds.), *Knowledge, Concepts, and Categories*. London: UCL Press.

Medin, D. L. & Wattenmaker, W. (1987). Category Cohesiveness, Theories, and Cognitive Archaeology. *In:* Neisser, U. (Ed.), *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization.* Cambridge: Cambridge University Press.

Shepard, R. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*, 125.

Tversky, A. (1977). Features of Similarity. *Psychological Review*, 84, 327—352.