COGNITION

Cognition 87 (2003) 1–32

www.elsevier.com/locate/cognit

Similarity as transformation

Ulrike Hahn[a,*], Nick Chater[b], Lucy B. Richardson[a]

[a]*Cardiff University, Cardiff, UK*
[b]*Warwick University, Coventry, UK*

Received 9 August 2001; received in revised form 3 July 2002; accepted 14 September 2002

Abstract

We propose that similarity is determined by the transformation distance between representations: entities which are perceived to be similar have representations which are readily transformed into one another, whereas transforming between dissimilar entities requires many transformations. We present three experiments that indicate that similarity is strongly influenced by transformation distance. These data present a challenge for featural or spatial accounts of similarity. Finally, we introduce a family of transformation-based accounts of similarity, called 'Representational Distortion', as a specific example of a transformational approach to similarity. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Similarity; Transformation distance; Representations

1. Introduction

The breadth of cognitive and social contexts in which similarity is invoked as an explanatory construct is vast. Similarity forms part of explanations of memory retrieval (Hintzmann, 1986), categorization (e.g. Hampton, 1995; Nosofsky, 1986), visual search (Duncan & Humphreys, 1992), problem solving (Gick & Holyoak, 1980; Holyoak & Koh, 1987), learning (e.g. Gentner, 1989; Ross, 1984), linguistic knowledge (e.g. Bailey & Hahn, 2001; Hahn & Nakisa, 2000) and processing (e.g. Luce, 1986), reasoning (e.g. Rips, 1975), as well as social judgement (Smith & Zarate, 1992).

However, similarity has not been without its detractors (e.g. Goodman, 1972; Murphy & Medin, 1985) who have criticized the notion as too vague (although see, for example, Goldstone, 1994; Hahn & Ramscar, 2001a). Such criticism is best countered through the success of accounts of empirical phenomena that are based on explicit measures of simi-

* Corresponding author. School of Psychology, Cardiff University, P.O. Box 901, Cardiff CF11 3YG, UK. Fax: +44-29-20874858.
E-mail address: hahnu@cardiff.ac.uk (U. Hahn).

doi:10.1016/S0010-0277(02)00184-1

larity. Current work in the context of categorization, for instance, has yielded a wealth of both explicit and empirically successful similarity-based accounts (see, for example, the collection by Hahn & Ramscar, 2001b).

The two classical approaches to theorizing about, and measuring, similarity are the spatial account (Shepard, 1957), which represents similarity in terms of distance in a psychological space, and the Tversky (1977) contrast model which views similarity as a function of common and distinctive features of the entities under comparison. Both of these accounts have successfully been used in cognitive modelling (see Ortony, 1979; Osherson, 1990; Tversky, 1977, for applications of the contrast model; for some of the many applications of spatial models see, for example, Nosofsky, 1986; Rips, 1975; Shepard, 1987).

Both accounts suffer, however, from a fundamental limitation (see Hahn & Chater, 1997, 1998a,b for more detailed discussion). They are restricted in scope by the fact that they define similarity over very specific – and very simple – kinds of representation: points in space or feature sets. But almost all theories of the representation of natural objects, from faces to auditory scenes, from visual textures to sentences, assume that these cannot be represented in line with these restrictions. Instead, representing such complex stimuli is typically assumed to require *structured* representations, that can explicitly describe objects and the relations between them, and how a whole can be decomposed into its component parts, its parts into sub-parts, and so on. Thus, specifying the structure of a physical object, such as a familiar animal or a human face, requires not just specifying what features it has, but, crucially, how they are interrelated. Similarly, specifying the structure of a sentence or a musical phrase requires specifying relationships between different parts (words, notes). Structured descriptions can express such complexity – mere lists of features or dimensional values cannot (Biederman, 1995; Fodor, 1975, 1999; Fodor & McLaughlin, 1990; Fodor & Pylyshyn, 1988; Hahn & Chater, 1998b; Marr, 1982).

The issue of the role of structure in similarity has been taken up by more recent structural alignment approaches to similarity (see Markman, 2001 for an overview) which developed out of research on analogy (Gentner, 1989). The present paper considers a recent theoretical approach to similarity, Representational Distortion (henceforth, RD; Chater & Hahn, 1997; Hahn & Chater, 1997), which aims to provide a theoretical framework applicable to similarity judgements over representations of arbitrary form, including structured representations. According to RD, the similarity between two entities is a function of the "complexity" required to "distort" or "transform" the representation of one into the representation of the other. The simpler the transformation distorting one representation to the other, the more similar they are assumed to be.

How can the complexity of the transformation between two representations be measured? At a theoretical level, Chater and Hahn draw on a branch of mathematics, Kolmogorov complexity theory (Li & Vitányi, 1997), that provides a rigorous and general way of measuring the complexity of representations, and transformations between them. In intuitive terms, according to Kolmogorov complexity theory, the complexity of a representation, x, is the length of the shortest computer program that can generate that representation. The length of this program is written $K(x)$. The idea is that representations that can be generated by a short program are simple; those that require longer programs are complex. So, a representation consisting of a billion 1s, although itself very long, can be generated by a short program, and therefore has low Kolmogorov complexity. On the other hand, the text of Hamlet, although very much shorter than one billion characters, cannot be generated by any simple program, and hence is much more complex. Random strings of symbols, finally, have the greatest complexity – the shortest program for a random string is merely a verbatim quote of that string (whereas, at minimum, Hamlet can be put through a text compression algorithm, taking advantage of the redundancies of English text). We shall not consider the virtues of this measure of complexity here, except to note that it supports substantial applications in the cognitive and computing sciences (Chater, 1996, 1999; Quinlan & Rivest, 1989; Rissanen, 1989; Wallace & Boulton, 1968).[1]

Kolmogorov complexity has a natural application as a measure of similarity between representations, A and B. The simplest measure is the length of the shortest program which takes A as input, and produces B as output – that is, the length of the shortest program that "distorts" one representation into the other. This quantity is called the *conditional* Kolmogorov complexity, and is written $K(B \mid A)$. This and related notions have been explored mathematically, with an interest in providing a notion of similarity that is relevant to reasoning by similarity in computational applications, such as machine learning (Bennett, Gács, Li, Vitányi, & Zurek, 1998; Li, Li, Ma, & Vitányi, 2001; Li & Vitányi, 1997).

According to this viewpoint, the degree to which two representations are similar is determined by how many instructions must be followed to transform one into the other. For example, the conditional Kolmogorov complexity between the sequence 1 2 3 4 5 6 7 8 and 2 3 4 5 6 7 8 9 is small, because the simple instructions *add 1 to each digit* and *subtract 1 from each digit* suffice to transform one into the other. In the same way, 1 2 3 4 5 6 7 8 and 2 4 6 8 10 12 14 16 (*multiply/divide each digit by two*) are presumed to be similar. On the other hand, 1 2 3 4 5 6 7 8 and 3 5 7 9 11 13 15 17 are viewed as less similar, because two operations are required to transform one into the other (e.g. *multiply by two and add one*). Finally, two entirely unrelated representations will be maximally dissimilar – because there will be no efficient way of transforming one into the other. Indeed, in this case, the most efficient transformation will simply involve deleting the first representation, and reconstructing the second from scratch – because there is no shared information between the objects that can be exploited. Note that structured representations, which cause difficulties for the spatial and featural views of similarity, can naturally be dealt with in the present approach. Tree structures (e.g. in a semantic network) can be transformed by transformational operations on trees; sentences can be transformed by linguistic operations at the phonological, morphological, syntactic or semantic levels, and so on.

Like the Tversky (1977) contrast model, RD encompasses a broad family of measures based on Kolmogorov complexities, including, for example, both asymmetrical and

[1] The emphasis on programs may suggest that conventional programming languages, such as C and Java, are in play – where these languages presumably have no cognitive relevance. But the mathematical theory is, importantly, neutral about the nature of the programming language in play, so long as they have a certain very minimal computational power (specifically, they must be 'universal' programming languages). Indeed, a key result of Kolmogorov complexity theory, the invariance theorem, is that the Kolmogorov complexity of a representation is invariant, up to a constant additive factor, with respect to the choice of programming language (Li & Vitányi, 1997).

symmetrical variants.[2] All of these, however, share the central property that similarity is a decreasing function of transformational complexity. Also among this class are featural and spatial models; RD can capture such accounts as special cases (see Chater & Hahn, 1997 for derivations) and hence does not contrast but rather subsumes these types of account.

RD should thus be viewed as a general framework for understanding similarity, not as a specific psychological model. In Marr's terms (Marr, 1982) RD theory seeks to characterize the computational level problem involved in determining similarity. At this level of generality, RD theory tries to provide a framework for understanding similarity-based processes in the spirit of rational analysis (e.g. Anderson, 1990; Oaksford & Chater, 1998), providing a derivation of the Shepard (1987) Universal Law of Generalization (Chater & Vitányi, 2002; Chater, Vitányi, & Stewart, in press). It also aims to provide an explanation for the utility of similarity in inference – why it makes sense, for example, to categorize items on the basis of the categories of similar items (Chater & Hahn, 1997).

Building a concrete psychological account of similarity would require a much more detailed approach (see Chater & Hahn, 1997 for some first steps). Ideally, we need to take account of (I) the nature of the mental representations that are relevant to making a similarity judgement, (II) the set of transformations or instructions that can be used to distort one representation into another, and (III) any constraints on the ability of the cognitive system to discover simple transformations between mental representations. Cognitive science is not well-developed enough, at present, to provide detailed constraints on the nature of mental representation and mental transformation (although see Palmer, 1982, 1983, 1991; Shepard, 1970). We stress, however, that, in this regard, the RD theory of similarity is in the same position as spatial and feature-based views. In the context of the spatial view, note that the structure and even existence of putative mental spaces, and the locations of representations of objects within these putative spaces is a matter of conjecture; similarly, in the context of the featural view, it is unknown whether items are represented in terms of sets of features, let alone what features are part of these putative representations. The present lack of a sufficiently well-specified general theory of mental representation necessitates a compromise in empirical work. One can focus either on fundamental, general predictions of an account or try to seek out successful operationalizations of its basic aspects such that specific predictions can nevertheless be made. Much

[2] Generally, we will want similarity to be symmetrical or nearly symmetrical (see, for example, Ashby, Maddox, & Lee, 1994). In Kolmogorov complexity theory, the most popular measures of 'information distance' are symmetrical – the simplest, max-distance, measures the distance between a pair of items by the largest of the Kolmogorov complexities, in either direction (Li & Vitányi, 1997). Another natural measure, sum-distance, simply sums the Kolmogorov complexity in both directions – it turns out that this, and other measures, are closely related, and there are some technical reasons to prefer max-distance (although these reasons may not be of relevance in a psychological context). But this still leaves a residual problem – that more complex objects are, other things being equal, more 'distant' than simple objects, because the relevant Kolmogorov complexities will necessarily be larger. This seems counter-intuitive – it seems, instead, that we should 'normalize' by the overall complexity of the objects concerned. This can be done in a number of ways, the technicalities of which are beyond the scope of this paper (Bennett et al., 1998; Li & Vitányi, 1997). Note that this normalization is also used in machine learning contexts, such as where Kolmogorov complexity-based methods are used to assess the similarity of sequences of DNA (Li et al., 2001). Specifically, Li et al. (2001) normalize max-distance by the Kolmogorov complexity of the most complex of the two objects being compared. We thank an anonymous reviewer for highlighting the importance of these issues.

of the ingenuity displayed by Tversky's research on the contrast model exemplifies the former strategy, whereas the widespread use of multi-dimensional scaling to supplement spatial accounts might be seen as an example of the latter.

Our aim in this paper is to provide evidence for the general class of transformational approach to similarity by concentrating on its most fundamental prediction: that the similarity between two objects is determined by transformational relationships. To the extent that this work provides evidence that a transformational account is viable, we hope that this will encourage the development both of RD, and alternative general models of similarity based on transformations, and subsequent empirical work to test between these. But here we concentrate on the logically prior question of assessing the empirical usefulness of the transformational approach considered quite broadly. The experiments we conducted for this purpose present a difficult challenge for classical theories of similarity and raise novel issues for any future work on similarity.

1.1. Previous work

The idea under experimental test in this article, that similarity is determined by transformations, seems so intuitive that one might expect that it would have numerous precursors in the literature. In a broad sense, the literature is full of such precursors – there is a very large body of research that concerns transformations, particularly in perception, which might reasonably be interpreted as describing why items are viewed as visually similar. But in a narrow and focused sense, precursors are surprisingly few: the considerable bodies of research on similarity and on perceptual transformations have been brought together only rarely. Here we briefly describe relevant previous work from three fields, language, perception and analogical reasoning – in each case, previous research, though suggestive, has not forged a clear connection between transformations and similarity. We then consider the small number of previous studies that do address directly the concerns of the present article.

The psychology of language might seem an obvious place to look for relationships between transformations and similarity, because linguistic theory provides a well-specified inventory of linguistic operations (including transformations) that might be used to modify one linguistic structure into another. Indeed, psycholinguistic studies have repeatedly made use of so-called edit distance measures to determine the lexical "neighbours" of a word which are thought to influence the word's linguistic processing (see, for example, Luce, 1986). The standard edit distance measure classifies as a neighbour each word which can be derived from the target word through the insertion, deletion or substitution of a phoneme (component sound), thus rendering "spat", "at" and "pot" as neighbours of the word "pat". Edit distance can be viewed as a measure of word similarity. However, it is widely acknowledged to be a crude measure and there is virtually no empirical work which has made any attempt to evaluate edit distance-based measures as measures of similarity, because word similarity has only been of incidental interest to past psycholinguistic research, rather than being viewed as an issue in its own right (but see Bailey & Hahn, 2001).

In perception, there has been considerable theoretical and empirical work using a transformational framework, but, again, with little direct focus on similarity. A range of theoretical frameworks have been proposed which view the process of making sense of the perceptual world as a process of uncovering sequences of transformations. Most fundamentally, the mathematical study of geometry, the mathematical framework in which much theorizing in perception is most naturally expressed, is fundamentally concerned with transformations. Specifically, the essential properties of a given geometry are typically represented by the group of transformations that leave those properties invariant, a view which originated with Klein's Erlangen Program in 1872.[3] Indeed, Klein, and later mathematicians (e.g. Weyl, 1952) believed that transformations have deep connections with perception, particularly in relation to aesthetic judgement. Furthermore, transformations have been viewed as a crucial tool in building models of the kinds of patterns occurring in the natural world – for example, in the development of Pattern Theory (e.g. Grenander, 1996) and its application to computer image processing and visual perception (e.g. Mumford, 1996). Transformations have, moreover, been important in a vast range of theoretical viewpoints in the study of perception – they can be viewed as governing the kinematics of the natural environment (e.g. Shepard, 1984), providing a 'derivation' of the structure of the perceived object (e.g. Leyton, 1986), or governing how the cognitive system generalizes from past examples (e.g. Feldman, 1996). Transformations, of a range of specific families, have also been central to a range of approaches to machine vision and pattern recognition, as well as computational theorizing about the human visual system (e.g. Basri, Costa, Geiger, & Jacobs, 1998; Hildreth, 1983; Ullman, 1996).

This theoretical perspective is complemented by a large body of empirical studies examining specific transformations, ranging from translations and mirror symmetry (e.g. Corballis & Roldan, 1974), to mental rotation (e.g. Cooper, 1976; McBeath & Shepard, 1989; Shepard & Metzler, 1971), to the detection of Glass patterns (e.g. Dakin, 1997), and, more recently, to complex operations such as morphing human faces (e.g. Beale & Keil, 1995; Young et al., 1997). Mental rotation experiments, for example, have found the degree of rotation of an object to predict recognition times. This evidence might be viewed as consonant with the present account, to the extent that the time needed to recognize the relationship between two objects is assumed to be a measure of their (dis)similarity. Such a link seems at the very least plausible given that there are independently motivated accounts relating recognition to exemplar similarity such as those of Nosofsky (1986) or Hintzmann (1986). So with this additional assumption, we could view mental rotation results as indicating that the less the (rotational) transformation between two objects, the more similar they are. Nonetheless, this line of argument provides at best indirect support for viewing similarity itself in terms of transformations, and it has not been discussed in the literature. A further limitation of object or face recognition studies as sources of evidence for similarity as transformation stems from the fact that they typically involve only a manipulation of degree within a single continuous transformation. A general theory of similarity founded on the notion of transformation will necessarily require not only evidence regarding a much wider set of transformations, but also data about the impact on similarity of combining multiple transformations of different kinds.

A substantial further body of previous research relevant to these latter questions is found

[3] We thank an anonymous reviewer for this point.

in the perception literature on figural regularity or "goodness" (e.g. Garner & Clement, 1963; Palmer, 1983) as well as occurring in theories of object recognition (Ullman, 1989, 1996). According to transformational accounts of goodness, for example, patterns are good insofar as there are many transformations under which the pattern is preserved (e.g. the number of axes of mirror or rotational symmetry). Thus, a circle (which is mirror symmetrical along any axis passing through its centre) is "better" than an ellipse (which is mirror symmetrical along just two axes). And goodness is not merely an intuitive notion – good patterns are recognized more rapidly, and perceived better in perceptually degraded conditions, than less good patterns. Note, though, that this literature is concerned with within-pattern relationships, rather than relationships between distinct patterns; this contrasts with the present focus on similarity, where the relationship of interest is between two distinct objects. Indeed, the picture is complicated further by the observation that the perception of certain symmetries appears to proceed differently, depending on whether the symmetry occurs within a pattern perceived as a single object, or whether it relates two patterns that are perceived as perceptually distinct objects (Olivers & van der Helm, 1998). The literature on perceptual transformations is, nonetheless, potentially extremely useful in constraining the set of mathematically possible transformations to those which may be carried out by the cognitive system. It is from this set that we recruit several of the basic geometric transformations such as translation, rotation, mirror image, dilation, and so on, which we manipulated in the experiments reported below.

Research on analogical reasoning has also involved discussion of transformations. Most relevantly, computational models of analogical inference frequently postulate that analogies are made between pairs of items by representing these items so that a relatively simple set of transformations can map one item onto the other. Examples of such transformations are geometrical transformations, as discussed in the perception literature, where analogies are between geometrical objects (Indurkhya, 1989; O'Hara, 1992). It also seems that analogical mappings and correspondences for a range of other materials, from strings of symbols, tabletop layouts, or scientific theories (e.g. Gentner, 1983; Hofstadter, 1997; Mitchell, 1993; Mitchell & Hofstadter, 1990), could readily be recast in transformational terms, regardless of whether the notion of transformation has figured in their explication. Developing a relationship between this research and the RD approach to similarity seems an interesting challenge for future research. It requires adopting the controversial stance that similarity and analogy are sufficiently close as to be modelled in a single framework (cf. Gentner, 1989); it further requires establishing the simplicity of a transformation (or sequence of transformations) as a measure of the strength of an analogy. One attraction of such an approach is that, given the close relationships between Kolmogorov complexity theory and theories of inductive inference, it might be possible to clarify *when* and *why* analogical reasoning is justified. This is a problem that has been extensively but inconclusively discussed in the philosophy of science (Achinstein, 1963; Hesse, 1964); a program of understanding similarity and analogy in terms of simple transformations has not, as far as we know, been developed within this literature.

In summary, despite considerable interest in the idea of mental transformations, in a range of fields, there has been relatively little research relating transformations directly to similarity. There are, however, at least three interesting exceptions, of direct relevance to the present paper. The first of these is a study by Posner, Goldsmith, and Welton (1967).

Participants were asked to rate the perceived distance of the distorted patterns to the original patterns, and perceived distance from the original was found to be a logarithmic function of the mean objective distance by which component dots had been moved between original and distortion.[4] This experiment provides a first test of the notion of similarity as transformation as it links transformation distance directly to estimates of perceived similarity. However, this result does not provide strong support specifically for the transformation-based account, because a spatial account of similarity can readily provide an equally good account of the data.[5] Below, we use stimuli that are designed to help differentiate between the transformational and competing accounts.

Though also not aimed directly at the issue of how similarity might best be characterized, a subsequent study by Franks and Bransford (1975) provides more general, if tentative, support for the notion of similarity as based on transformation distance. Franks and Bransford sought to extend the work of Posner and colleagues, particularly on prototype abstraction (e.g. Posner & Keele, 1970), and replaced the original dot patterns with simple geometric figures. Underlying Franks and Bransford's stimulus set was a prototype, which was not shown during training. All other items in the stimulus set were derived from this prototype through the application of one or more simple transformations such as deletion, substitution or permutation of component parts, and it is this use of a broader set of transformations that makes the study especially relevant from the current perspective. Participants were exposed to a subset of the items and were then given a recognition test coupled with confidence ratings. The set of test items consisted of the prototype and a series of novel items which varied in their transformational distance to the prototype (i.e. in the number of operations necessary to generate the item from the prototype). Recognition ratings were directly related to transformational distance to the prototype, with the prototype itself receiving the highest rating. Franks and Bransford viewed this as support for Posner's claims about prototype abstraction. As in the case of the mental rotation experiments or face recognition results cited above, these findings might be taken as indirect support for the notion that similarity is related to transformational distance if it turns out to be the case that similarity dominates recognition time, thus allowing the reverse inference from recognition times to between-item similarities in general (e.g. Nosofsky, 1986).

Finally, Imai (1977) proposed that pattern similarity between strings of either filled or unfilled circles was based on transformational structure. He found support for this claim in terms of a qualitative relationship between the number of transformations between two

---

[4] Interestingly, Posner, Goldsmith, and Welton (1967) created distortions of their stimuli using a specific probability distribution, of distortions around the original pattern, and measured the magnitude of the distortion by an entropy-based measure, related to the negative logarithm of the probability of the new stimulus arising as a distortion of the original. This kind of information-theoretic measure has some interesting technical relationships (though it is not identical) with the mathematics of Kolmogorov complexity used in the RD account of similarity. See Li and Vitányi (1997) and Chater (1996) for discussion.

[5] Specifically, the judged distance between patterns was proportional to the average distance between the component dots of the pair under comparison (see Posner et al., 1967, Experiment 2). This average distance is itself proportional to the distance (using a city-block metric) between a spatial representation, in which the position of patterns in the n co-ordinates of a pattern in 2D space is represented in a $2n$-dimensional space (formally, this is a product space of the 2D spaces associated with the locations of each individual point).

patterns including insertions, deletions, mirror imaging and reversals (Imai's transformations are described in more detail below), and their judged similarity. However, no statistical analysis was performed, nor is this possible with the data as reported.

In conclusion, there is currently no clear-cut experimental evidence for the centrality of transformations to a general theory of similarity, despite a wealth of relevant research on cognitive transformations and on similarity. Much research has been guided by the intuition that similarity is affected by the degree of variation along a particular transformation. More importantly there is also some work, such as Franks and Bransford's studies on prototype abstraction, Imai's demonstrations on similarity, and psycholinguistic studies drawing on edit distance, which have been guided by an implicit intuition that the number of different transformations combined affects similarity. All of these studies provide data that are encouraging. However, none of these studies have had the goal of providing a general theory of similarity.

In consequence, these studies do not motivate, state or test with any generality the hypothesis that similarity is based on degree of transformation. What is at issue in the present paper is not simply whether particular transformations affect degree of similarity as might be inferred from these previous works, but rather whether similarity itself is a notion best explained in terms of transformation distance. That manipulations of particular transformations have widely been shown to influence similarity is encouraging for this latter view. But the view that similarity, in general, is a function of transformation distance has not yet been adequately tested.

1.2. Experimental investigation

As outlined in Section 1, our goal is to seek empirical support for a general transformational approach to similarity, not – at this point in time – to test specific properties of the RD family of models. Specifically, we seek to provide evidence that the general notion of transformation should form the central building block of a theory of similarity. To this end, we aim to evaluate the empirical usefulness of a general transformational perspective on similarity on two key criteria. The first criterion concerns whether a general transformational perspective allows better coverage than current featural and spatial views of similarity, i.e. can it account for similarity judgements for stimuli to which these accounts do not apply successfully? As will be discussed in more detail below, featural and spatial models of similarity allow only a very restricted set of possible transformations between two objects. Consequently, this first criterion can be met by experimental evidence that shows similarity judgements to be influenced by transformations outside of this restricted set. Such evidence will indicate the need for a new approach to similarity.

More is required, however, to make plausible the claim that similarity should be conceptualized in terms of the general notion of transformation. It must be shown that the notion of transformation itself provides the right level of abstraction at which to seek the systematic relationships governing the similarity between objects. This is the second criterion. It can be met by the empirical demonstration that similarity seems to be influenced by a wealth of specific transformations that share nothing apart from the fact that they are all transformations.

That both these criteria can be met is crucial if similarity is plausibly to be thought of as determined by transformational relationships. There would be no point in pursuing the project of a transformationally-based general theory of similarity any further if these two criteria could not be fulfilled; they can be viewed as logical preconditions for a transformational view of similarity and so constitute the obvious experimental starting point.

Because they focus only on the most general core of the class of transformational accounts, meeting these criteria will not suffice to establish RD in preference to other possible transformational theories. However, they suffice to test whether or not there is any point in embarking on the laborious path of developing and testing particular transformational accounts. At the same time, their generality has the benefit that the relevance of results – if obtained – will extend beyond the confines of any narrow technical debate.

To test the similarity-as-transformation hypothesis according to these criteria, we require experimental data that directly link manipulations of transformation distance to perceived similarity. Moreover, we need to do this with a range of different kinds of experimental materials, chosen so that the transformational account can be differentiated from featural and spatial accounts as more restrictive, alternative conceptualizations of similarity and so that transformation can be established as the appropriate level of abstraction. We conducted three experiments to test transformational distance as a predictor of perceived similarity in this way. Though the basic design and procedure of all three experiments are the same, they make use of very different stimulus materials and transformations.

2. Experiment 1

An ultimate understanding of similarity as transformation will encompass a clear understanding of the set of relevant transformations involved. Undoubtedly, there are many possible transformations between two representations that have no psychological relevance, because the cognitive system does not use these transformations. Although the perception literature, as mentioned above, provides suggestions concerning what transformations are psychologically relevant, any inventory of transformations is inevitably tentative – just as an inventory of specific features or dimensions is inevitably tentative for featural and spatial views of similarity. In this situation, the best we can do is pick a set of transformations and combinations of transformations which we hope will be psychologically real, devise stimulus materials which manipulate these transformations, and see whether the predictions associated with these materials are born out. On the positive side, we must commit to a set of transformations in constructing the stimuli; hence the transformations do not provide degrees of freedom that we can exploit in the light of our results.

The most natural starting point for a first study was a set of transformations based closely on previous work. Experiment 1 was based directly on the work of Imai (1977) with sequences of filled or unfilled circles as this was the most direct investigation of transformations and similarity to date. The basic transformations between two items in our materials, as in Imai's materials, were mirror imaging, reversal, phase shift, insertion and deletion, which were necessary to convert the test stimulus into the target. An example stimulus pair from Experiment 1, shown in Fig. 1, illustrates one of these basic transfor-

Fig. 1. Sample stimuli from Experiment 1.

mations, a "phase shift", whereby the two rows of "blob" patterns can be transformed into one another through the application of a single operation which uniformly shifts all blobs in the sequence one position to the left.

Transformation distance was operationalized as the number of basic transformations of this kind which had to be combined in order to transform one item into the other. In other words, two items which could be transformed into each other through the application of a single transformation such as a phase shift were assumed to be more similar than items which required, for example, a phase shift plus the deletion of a component blob (for examples see Table 1 which contains all comparison pairs and their assumed transformational relationship). The transformational account is then straightforwardly tested by examining how well the transformation distance of each pair of items (i.e. the number of transformations linking them) predicted the pairs' perceived similarity.[6]

How can we additionally derive evidence against featural accounts from these materials? As indicated above, the notion of similarity as transformation encompasses featural views of similarity as special cases. On any featural story, similarity is a function of matching and/or mismatching features (e.g. Tversky, 1977). From a transformational perspective, such featural models simply construe similarity between objects as the result of an extremely limited set of transformations, namely feature insertion, feature deletion, or feature substitution (likewise spatial models whose transformations are restricted to increments along continuous valued dimensions; see also Chater & Hahn, 1997). From a transformational perspective, such models are consequently not "wrong" as such, but merely too restrictive. Providing evidence against featural models thus requires showing that these restrictions limit their ability to explain performance. The way we seek to do this in Experiment 1 is by directly comparing the predictive accuracy of our measure of transformation distance and a featural distance measure. The "features" of objects are notoriously difficult to pin down, given that seemingly any property of an object, whether this be a simple attribute such as "colour" or a higher-order property such as "symmetry", can be deemed a feature (and indeed, at least to Tversky, 1977, this has been presented as an advantage of such models). This complicates any evaluation of featural theories, but our stimulus materials support naturally the assumption that each of its component blobs constitute a "feature" of the row in question. The number of mismatching component blobs between two patterns thus provides a straightforward featural indicator of their similarity. This metric can be compared with our transformation distance measure despite the fact that feature insertions, deletions and substitutions are themselves basic transfor-

[6] We note that it is likely that these transformations themselves differ in complexity, in the representational language of the cognitive system, and hence that a weighted sum, rather than an unweighted sum of the transformations would provide a better fit with the data. Lacking independent means of determining the relative complexity of these transformations, however, we decided not to allow ourselves the additional degrees of freedom implied by weighting. More generally, number of transformations serves as a relatively crude approximation for the code length for a sequence of transformations, that underlies similarity according to the RD account, as discussed later in the paper.

Table 1
Stimuli used in Experiment 1

No. of transformations	Type	Stimuli: Item one	Stimuli: Item two
1	Reversal	●○○○○○●●○●●●	○●●●●●○○●○○○
		●●●○●○●○●○●●	○○○●○●○●○●○○
		●●●●○○●●●●○○	○○○○●●○○○○●●
		○○●●○●○●○○○○	●●○○●○●○●●●●
1	Mirror	○●●●○○○●○●●○	○●●○●○○○●●●○
		○○○●○○○●●○●●	●●○●●○○○●○○○
		●●○○●●●○○○○○	○○○○○●●●○○●●
		●○○●○○●○○○●○	○●○○○●○○●○○●
1	Phasic	●●○○●●○●○●○●	●○○●●○●○●○●○
		●●●●○○○○●●●●	○●●●●○○○○●●●
		●○○○○○○○●●●●	●●○○○○○○○●●●
		●○○●●○○●●○○●	●●○○●●○○●●○○
1	Deletion	●●○●●●○●●○●○	○●●●○●●○●○
		●●●○○●○●○●●●	●●○○●○●○●●●
		○●○●○●○●○●○●	○●○●○●○●○●
		○○○○●●●●○○○○	○○○○●●●○○○○
2	Reversal & Mirror	●○○○●○●○●○○●	○●●○●○●○●●●○
		●○○○●●○○●○●●	○○●○●●○○●●●○
		●●●●○●●●○●○●	○●○●○○○●○○○○
		●●●○○●●●○●○●	○●○●○○○●●○○○
2	Deletion & Mirror	●●○○○●●○○○○●	●○○○○●●○○○
		○●●○●●○●○○●●	●●○○●○●●○●●
		●●○●○○●○●○○○	○○●○●○○●○●●
		○●●●●○○●●●●●	●●●○○●●●●○
2	Reversal & Phasic	●●○●●○●●○●●○	○●○○●○○●○○●○
		○○○●○○○●○○○●	●○●●●○●●●○●●
		○○○○○●○○○○○○	●●●●●●●●○●●●
		●○○●●●●●●●●●	○○○○○○○●●○○○
2	Insertion & Phasic	●●●●○○●●●●●	○○●●●●●○○●●●
		○●●○○●○●○●●	●○●○●●○○●○●○
		●●○●●●●○●●	●●●○●●●●○●●●
		●●●○●●●○●●●	●●○●●●○●●●○●
3	Deletion, Reversal & Phasic	○○○●●○○○●●○○	○●●●○○●●●○○
		●○●○●●○○○●○●	●○●○●○○●●●○
		●○○○○●○○○○○●	●●○●●●●●○●●
		●●●○●●●○●●●○	○●○○○●○○○●○
3	Deletion, Reversal & Mirror	●○○●○●○○●○○○	●●●○●●○●○●●
		●○○●●●○●●○○●	○●●○○●○○○●●
		○○○○●●●●●●○●	○●○○○○○○●●
		○●○●○○○○●●●●	○○○○●●●●○●○
3	Insertion, Reversal & Phasic	●○○○●○●○○○●	○●○●●●○●○●●●
		●○○●○○●○○●	●○●●○●●○●●○●
		●●○○●●●●○○●	○○○○●●○○○○●●
		○○○●●○○○●●○	○○●●●○○●●●○○
3	Insertion, Reversal & Mirror	○●●●○○●●●●●	○○○○○●●○○○○●
		○○○●●○○○○●	○●●●●○○●●●○○
		●○●●○●○○●●	○○●●○●○○●○○●
		●○●○○●○●○○○	●●●○●○●●○●○○
Control	-----	●●●●●●●●●●●●	○●●○○○●○●●○●
		○○●○●○○○●○●○	○●○○○○○●○○●●
		○●●●●●○○○○○○	○●●○○●○○●○○●
		●●○●○●○●○●●●	○●●●●●○○○●●○
		●●●●●○○○●○●○	○●○○●●○○○●○○
		○○○●○○○●○○○●	●○●○○○○○●●●○
		○●●●●○●●●●○○	●○○●○●●○○●○●
		●●○○●○○●●○●●	○●●●●●●●○●○○

mations, because many of our basic transformations affect more than one component (and thus "feature") at the same time. In the Fig. 1 example of a phase shift the two rows of "blob" patterns are viewed as transformationally very similar because they can be transformed into one another through the application of a single operation which uniformly shifts all blobs one position to the left. From the featural perspective, however, patterns should be similar to the extent that they overlap, and hence a 'shift' operation between two items, which destroys correspondences between the two patterns, should be associated with low similarity. Similarly, mirror imaging and reversals constitute single transformations while affecting all component "blobs" in the string. We can thus use materials of this kind not only to test the prediction that two patterns will be perceived to be similar when they are transformationally close, but we can directly compare the transformational predictions with those of a featural account and test whether the number of transformations between the two patterns is a better predictor of perceived similarity than is the number of mismatching individual component blobs.

2.1. Method

2.1.1. Participants

The 35 participants were all first year undergraduate students taking an option in Psychology at the University of Wales, Cardiff, who took part in this experiment in return for entry into a draw to win one of several prizes of £10. They consisted of five males and 30 females, whose ages ranged between 18 and 46 years ($M = 20.9$ years, $SD = 5.78$ years).

2.1.2. Materials

The stimuli consisted of strings of black and white dots that were presented in pairs for participants to judge the level of similarity between the pair. The stimulus designs were based upon the work of Imai (1977) who identified four different basic transformations one could apply to a series of dots, namely mirror images (one string of dots is a reflection of the other), phase shifts (the pattern remains the same, but is shifted to the left or right along the string of dots), reversals (black dots become white and white dots are changed to black), and wave length changes (the series remains the same, but the magnitude of the pattern is adjusted, for example black-white-black-white becomes black-black-white-white-black-black-white-white). For the purposes of this study, the key transformations utilized were phase shifts, reversals and mirror images, with the addition of insertions (one or more consecutive dots added to a string) and deletions (one or more consecutive dots removed from the string). Any one of these operations in isolation constituted a single transformation. There were 16 examples of single transformations in this experiment (four each of phasic, reversal, mirror and deletion).

The multiple transformations were achieved by combining two or three of the above operations and performing them one at a time upon a string of dots until the final image was created. There were 16 examples of two-transformational changes (four reversal and mirror, four reversal and phasic, four deletion and mirror, and four insertion and phasic) and 16 examples of three-transformational changes (four deletion, reversal and mirror, four deletion, reversal and phasic, four insertion, reversal and mirror, and four insertion, reversal and phasic).

As a control, there were eight pairs of stimuli that were deemed to be unrelated or so multiply transformed as to make the transformations unperceivable. These were designed so that it would be simpler for a person wishing to instruct another to create one string from the other to start with a blank string rather than trying to adjust the target string of dots. Adding these control pairs to the one, two and three transformation items yielded a total set of 56 pairs of stimuli items. These can be viewed in their entirety in Table 1, together with a description of the transformations that constitute the relationship between each pair.

Although it is conceivable that another person could examine these items and disagree with the experimenter as to the number of transformations required to derive a particular image from its target, each individual stimulus item was actually created from the target using a specific quantity of operations so that this study is based upon a priori assertions as to the number of transformations required. Thus, we are committed to the number of transformations for each image in advance of any data collection and any ambiguities in the number of transformations will diminish our predictive accuracy.

Each pair of stimuli was assigned with a random number from 11 to 66 (single digit numbers were not used, in case participants mistakenly assumed that they referred to the similarity ratings), and printed horizontally, side by side (with mirror images always presented so that the invisible line of reflection would be between the two strings and not from the outer edges) onto a single sheet of paper together with brief instructions and a rating scale of 1 (very dissimilar) to 7 (very similar). These sheets were then placed into a different random order for each participant and bound into a booklet.

2.1.3. Procedure

Participants were instructed to complete the booklet in one sitting and to work through the items in order. The information within the booklet directed them to examine each pair of series of blobs and assess the similarity of their initial perception of the two strings of blobs. The participants were asked to indicate the similarity by circling a number on the rating scale of 1 to 7 presented on each page in conjunction with each pair.

2.2. Results

Bivariate correlations between number of transformations and mean similarity rating of each item were found to be highly significant with Spearman's rho $= -0.69$ ($P < 0.005$). The comparison featural model which left aligned the two rows of blobs and counted the number of (mis)matching features fared considerably worse: Spearman's rho $= -0.28$ ($P < 0.05$).

Analysis of individual subject ratings confirmed these findings, revealing great conformity across participants: 25 of 35 participants showed significant correlations as predicted. Such consistency was not found using the featural model, with only eight of the 35 participants showing a significant correlation.

The general relationship between number of transformations and mean similarity ratings is graphed in Fig. 2 which suggests, somewhat surprisingly (see, for example, Shepard, 1987), an approximately linear relationship.

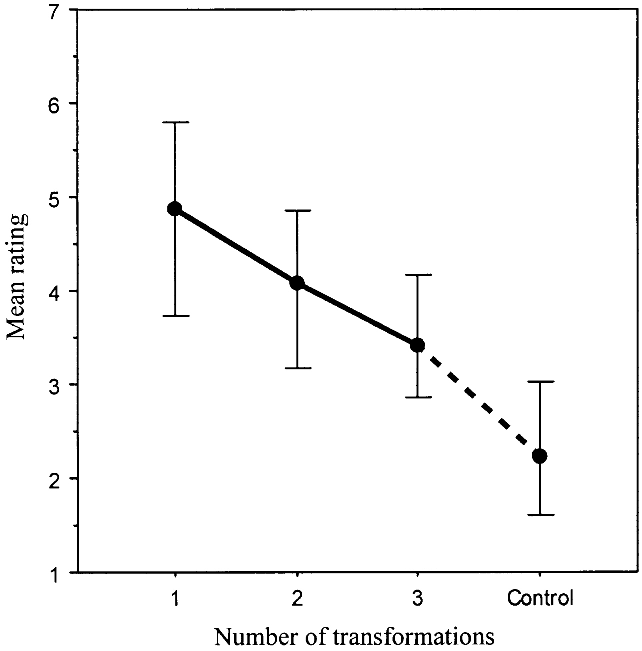


Fig. 2. Experiment 1 (circle series): graph depicting the relationship between transformation distance and mean similarity rating, from 1 (very dissimilar) to 7 (very similar). The error bars represent the minimum and maximum ratings assigned by participants.

2.3. Discussion

Experiment 1 provides evidence of a statistically significant relationship between transformation distance and perceived similarity to complement the qualitative analysis of Imai (1977). The results of Experiment 1 thus confirm both Imai's original intuition and the predictions of the RD framework that transformations are central to similarity. In comparison, the featural model fares considerably worse in predicting participants' ratings. Consequently, the results also provide evidence for the limitation of featural approaches more generally.

It is, of course, possible that more powerful featural descriptions of the data could be found, but, at present, none are available. Furthermore, while it seems plausible that there could be higher-order properties of the patterns relating to more than one component which a more sophisticated featural description of these materials might wish to assume, it seems hard to see how the individual component features which we assumed could not be *part* of such a featural description and, as a result, heavily influence similarity. Ultimately, however, any challenge to the result that transformations proved the better predictor would have to take the form of an explicitly formulated featural account that provided better data fits. Furthermore, any putative featural explanation of this kind requires an independent motivation of the features adopted, that is, the postulated features must themselves not be motivated exclusively by salient transformations. Otherwise, the

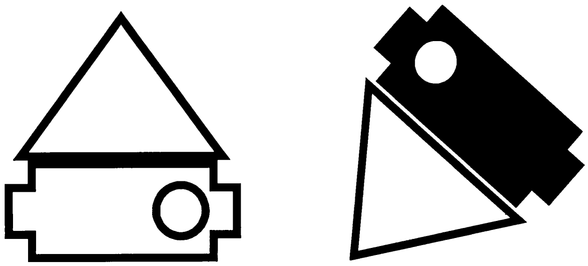


Fig. 3. Sample stimuli from Experiment 2.

featural description becomes an entirely redundant, ad hoc mimicry. The materials of Experiment 2 make this latter point more clearly.

3. Experiment 2

The second experiment used simple geometric shapes of the kind that has been popular in previous similarity research (e.g. Goldstone, Medin, & Gentner, 1991). The relationships between comparison items were, again, based on a range of different transformations. Our choice of relevant transformations was informed by the above mentioned perception literature. For example, the pair of items shown in Fig. 3 has a transformation distance of two as they can be made identical through rotation and filling in of the base.

For these materials, there is already no obvious way to calculate a featural model for contrast purposes. Furthermore, many of the "features" such as the orientation of an item in a pair where one has been rotated are salient only *because* of the relevant transformations. This is well-illustrated in an example by Goldstone (1994).[7]

When seen in isolation, the stimulus shown in Fig. 4a is most likely to be described simply as "a circle in a square". Only when contrasted with a similar image containing a filled circle within a square (Fig. 4b) does the fact that the circle is also white become salient. The contrasts available in the immediate comparison context influence our mental representations of the objects under comparison (see also Medin, Goldstone, & Gentner, 1993). For our stimuli this means that central object "features" will be derivative on the transformations present: for example, orientation is unlikely to have cognitive salience in a comparison *until* orientation is manipulated through rotations. Consequently, though it might, in principle, be possible to derive featural descriptions for our stimulus items, these descriptions would be likely to implicitly underscore the importance of transformations, rather than providing an alternative to relying on transformations.

As in Experiment 1, the prediction is that number of transformations will be negatively correlated with degree of perceived similarity.

[7] Keynote address to the Interdisciplinary Workshop on Similarity and Categorization, University of Edinburgh, 1997.

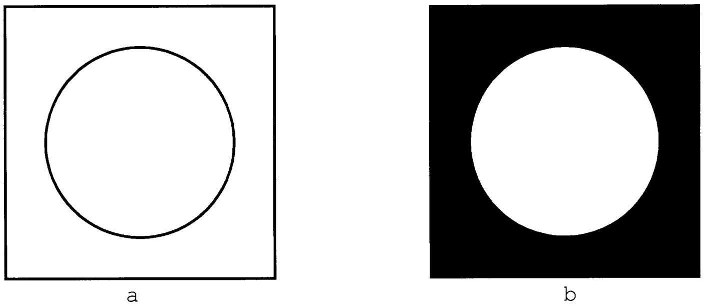


Fig. 4. Example of the role of contrast in determining featural descriptions: viewed in isolation the object a is typically described as "a circle in a square", in the context of object b, however, this becomes "a white circle in a white square".

3.1. Method

3.1.1. Participants

The 21 participants were all undergraduate students at the University of Warwick. They took part in this experiment, in combination with another study, in return for the payment of £5. They consisted of eight males, 12 females and one person who wished to remain anonymous. Their ages ranged between 18 and 37 years ($M = 21.7$ years, SD = 4.38 years).

3.1.2. Materials

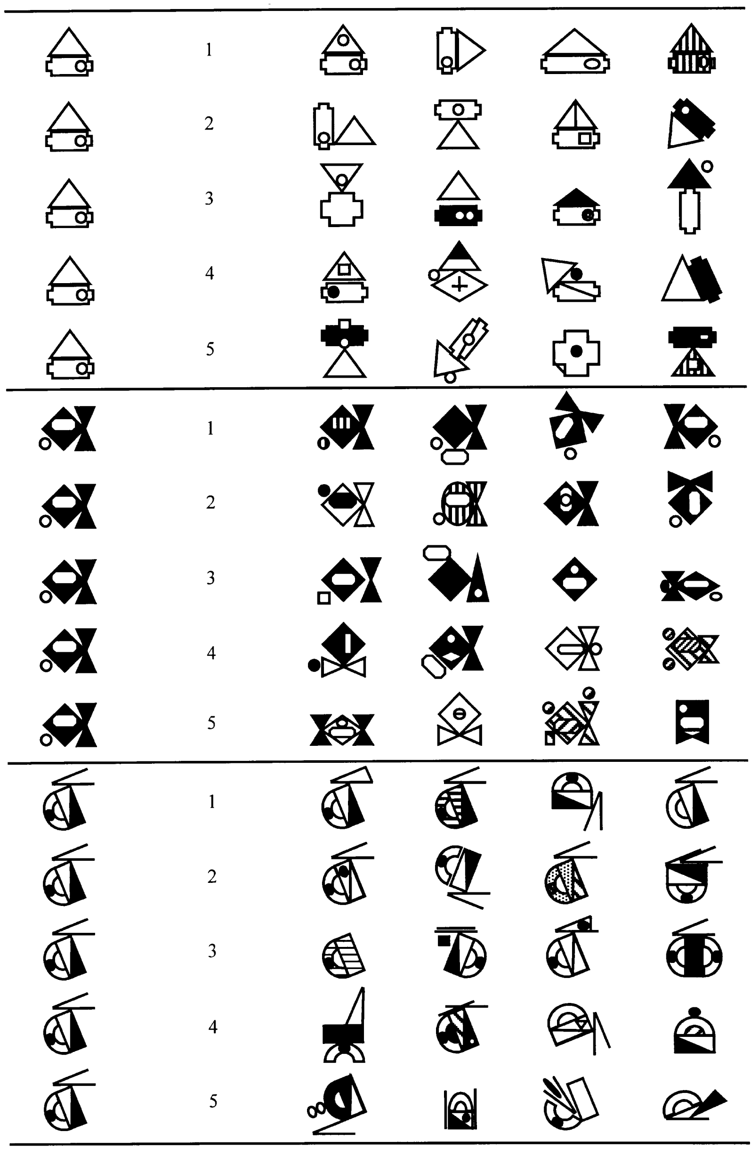
The stimuli consisted of line drawings produced with Microsoft Word. These drawings were created using a combination of "autoshapes" that were presented in a particular formation and with certain filling patterns to create a "target" image. To construct the stimuli set, the experimenter began with the target geometric shape and made adjustments to it, each of which constituted one transformation. Examples of single adjustments could be stretching the whole object, rotating the object, changing all filled-in black areas into stripped areas, moving one of the component shapes to a different location in relation to the rest of the image, reflecting all or part of the image, adding a new shape into the image or removing a component item. There were three separate target stimuli and for each of these targets four examples of one-transformation changes were created.

Multiple transformations were constructed for the same three target images by using a combination of the above techniques, which were applied to the image one at a time. There were four examples each of two, three, four, and five transformations. This made a set of 20 pictures for each base target and so 60 comparison pairs in total. All 60 images, together with their targets, are presented in Table 2.

Each transformed geometric shape was printed onto a separate page together with its corresponding target, which was always placed to the left of the transformed shape. These pages were assigned a random number from 11 to 70 (single numbers were not used to avoid participants confusing them with the rating scale). A set of instructions was at the top of each page and below them, a rating scale from 1 (very dissimilar) to 7 (very similar).

Table 2
Stimuli used in Experiment 2

Original stimulus	No. of transformations	Stimuli



3.1.3. Procedure

The participants were handed the booklet and told that this was a study on similarity and that their task was to examine each pair of pictures and assess how similar they perceived each pair to be to one another, based upon their initial impressions of the two items. Participants indicated this perception of similarity upon a 7-point rating scale that was printed on each page together with a repetition of the instructions. Their task was to circle the number that corresponded to their rating, from 1 (very dissimilar) to 7 (very similar). They worked through the booklet in the presence of the experimenter and completed each page in the order in which they were presented.

3.2. Results

Bivariate correlations between number of transformations and mean similarity rating of each item were highly significant with Spearman's rho $= -0.89$ ($P < 0.001$). Analysis of individual participants' ratings again revealed great consistency across subjects, with 19 of the 21 participants showing a significant correlation as predicted.

The general relationship between number of transformations and mean similarity ratings as graphed in Fig. 5 closely matches that found in Experiment 1. The relationship between transformation distance and similarity again is approximately linear.

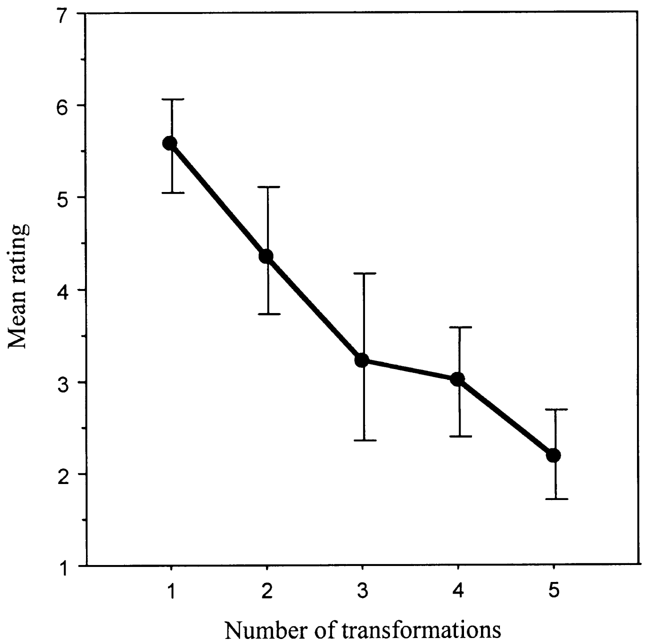


Fig. 5. Experiment 2 (geometric shapes): graph depicting the relationship between transformation distance and mean similarity rating, from 1 (very dissimilar) to 7 (very similar). The error bars represent the minimum and maximum ratings assigned by participants.

3.3. Discussion

The results of Experiment 2 provide further support for the role of transformations in the context of similarity. The results obtained mirror those of Experiment 1 almost exactly, despite the fact that the stimulus materials are very different.

The results also provide further evidence for the limitations of featural and spatial accounts, despite the fact that no explicit comparison models were tested, given the way the notion of transformation goes beyond featural or spatial accounts. The materials of Experiment 1 highlighted the way in which featural or spatial accounts are likely to remain at odds with transformational predictions for all those cases where a single transformation affects multiple aspects of an object (with phase shifts a prime example). The materials of Experiment 2 supplement this by illustrating how even natural "features" such as orientation are influenced by transformations. Many of the very "features" which a featural account might posit seem salient because of the transformational relationships that obtain between the two objects of the comparison. This is indicative of the general bidirectional relationship which RD theory posits between object representation and transformation, with perceived transformations influencing which aspects of an object become salient and vice versa (Hahn & Chater, 1998c; see also Hofstadter, 1997; Hofstadter & Mitchell, 1994).

4. Experiment 3

The materials of Experiments 1 and 2 were amenable to basic featural or spatial representations and comparison processes, even if these are, as argued, empirically and conceptually less satisfactory than the explanation provided by transformation distance.

The stimulus materials of Experiment 3 sought to take the argument against featural and spatial representations one step further, by using materials for which such representation schemes appear especially inappropriate: namely, materials where similarity is determined primarily by relational information. As indicated in Section 1, the need to incorporate relational information, which appears to require structured representation, has previously been identified as a critical limitation of featural and spatial schemes. The issue has been addressed both theoretically in the context of critical examinations of connectionist networks (Fodor & McLaughlin, 1990; Fodor & Pylyshyn, 1988) and empirically in the context of structural alignment models of similarity (see Markman, 2001, for an overview). Experiment 3 seeks to add to this body of evidence by using intrinsically and necessarily relational materials. These take the form of simple three-dimensional objects which are assembled from (typically) three individual Lego bricks: one large brick, coloured blue; a medium size yellow brick; and a small red brick. A schematic sample illustration is provided in Fig. 6.

Each similarity comparison comprised two objects assembled from these three bricks, albeit in different spatial arrangements. Despite the extreme simplicity of these stimuli, relational information (i.e. information about relative position, such as, for example, that the red brick is on top of the yellow brick, etc.) is paramount to the representation of the composite objects. However, the interest of these materials is not limited to the difficulties

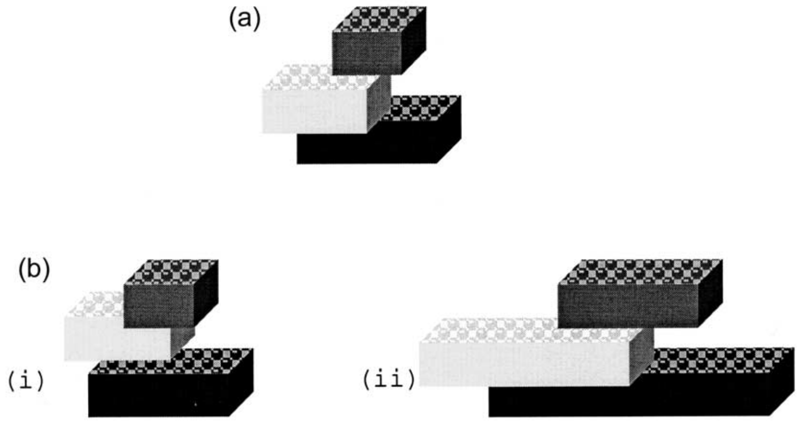


Fig. 6. (a) Experiment 3: the target Lego brick arrangement. (b) Experiment 3: two examples of single transformations from the target Lego brick array. In the first item (i) the yellow brick has been moved backwards by one notch, while the red brick remains in the same position relative to the blue brick. In the second item (ii) the bricks are in an identical arrangement to the target array, but the *whole* item has been "stretched" with the four-point red brick becoming an eight-point brick, the six-point yellow brick changing to 12 points and the eight-point blue brick becoming a 16-point brick.

they pose for featural or spatial accounts. From a transformational perspective, the Lego brick objects are of interest for two reasons. First, they allow an initial examination of the role of transformations in the similarity assessment of real-world objects, albeit maximally simple ones. Second, these materials support a whole new range of transformations to complement those investigated in Experiments 1 and 2. Our assumption, here, was that the judged similarity between pairs of objects would be determined primarily by the physical manipulations required to turn a target object into the comparison object.

4.1. Method

4.1.1. Participants

The 27 participants were all first year undergraduate Psychology students at the University of Wales, Cardiff, who took part in this series of experiments in return for course credit. They consisted of six males and 21 females, whose ages ranged between 18 and 32 years ($M = 19.9$ years, SD $= 3.11$ years).

4.1.2. Materials

The stimuli were based upon an initial "target" array of three Lego bricks (a four-point square red brick, upon a six-point oblong yellow brick, on top of an eight-point oblong blue brick) arranged into a particular three-dimensional structure (see Fig. 6).

Apart from the two examples that were chosen to be totally unrelated to the target, all of the Lego stimuli were constructed by transforming the original target object a set number of times. In each case, the experimenter began with the target arrangement and made adjustments to it, each of which was deemed to constitute one transformation. For example, one adjustment (or transformation) could involve moving an object within the arrangement while keeping the other bricks in the same relative positions to one another (see Fig. 6i). Another alternative would be to “stretch” a section of the arrangement or the entire structure (see Fig. 6ii). Other single transformations could involve substituting a brick for one of a different size or colour, adding an additional brick, subtracting an existing brick, or changing the order of the bricks within the arrangement.

As in Experiment 2, these materials might potentially be viewed as the result of different transformations than those used in their creation. For example, Fig. 6ii depicts a stimulus item created by stretching the whole arrangement, which is counted as one transformation. It could alternatively be classed as three separate transformations (each individual brick being stretched); again, our predicted transformational distance for each item is committed to in advance of data collection, so that any ambiguities of this kind will only decrease our predictive accuracy.

Performing a series of different single adjustments to the same array one after another created stimuli with multiple transformations. The transformations performed to create each arrangement were counted and the entire stimulus set was constructed so that it comprised ten items with one transformation, ten with two transformations and five items each with three, four, five and six transformations away from the “target”. The transformed stimuli, together with two examples of unrelated (multiply transformed) stimuli, made a total stimuli set of 42 items.

The unrelated stimuli were classed as such because starting with the initial target array would hold no benefit for the construction of these items. Unlike the “short-cuts” to construction which starting from the target array provides in the case of the transformationally related stimuli, the target, here, provides no clues as to the structure, shape or constituting pieces of the unrelated comparison item. It would be simpler to discard the target and begin creating the unrelated item from scratch.

Once these arrangements of Lego bricks had been constructed, each was glued into a permanent structure and a small label with a number from 1 to 42 (selected randomly, but not viewed by the participants) was applied to the underside of each stimulus item for recognition purposes.

Participants were given an A4 sheet of paper that clearly showed the rating scale from 1 to 7 with 1 being “very dissimilar” up to 7 being “very similar”, as a reference and reminder as to the scale.

4.1.3. Procedure

Participants were shown the “target” object made of Lego bricks and told that they would be presented with a series of other arrangements of Lego bricks. They were instructed to rate how similar their perception of each stimulus was to the target, on the scale of 1 (very dissimilar) to 7 (very similar) which was presented in front of them. The Lego stimuli were presented to each participant in a different random order; every participant rated all of the Lego stimuli within the set of 42 items.

4.2. Results

Bivariate correlations between number of transformations and mean similarity rating of each item were again highly significant with Spearman's rho $= -0.76$ ($P < 0.005$). Analysis of individual participants' ratings again revealed great consistency across subjects, with all 35 exhibiting a significant correlation between number of transformations and rated similarity. The general relationship between number of transformations and mean similarity ratings as graphed in Fig. 7 is again very similar to that that found in Experiments 1 and 2.

4.3. Discussion

Despite the very different materials and set of relevant transformations, the results of Experiment 3 closely match those of Experiments 1 and 2. Again, they provide evidence for the importance of transformations in explaining similarity judgements, and are difficult to account for in terms of the featural or spatial views of similarity, which cannot easily handle relational information. Thus, these results are in line with the predictions of a central tenet of the RD account: that the transformational relationships between representations of two objects determine their judged similarity.

But the results of Experiment 3 also have broader implications. The inherent relational

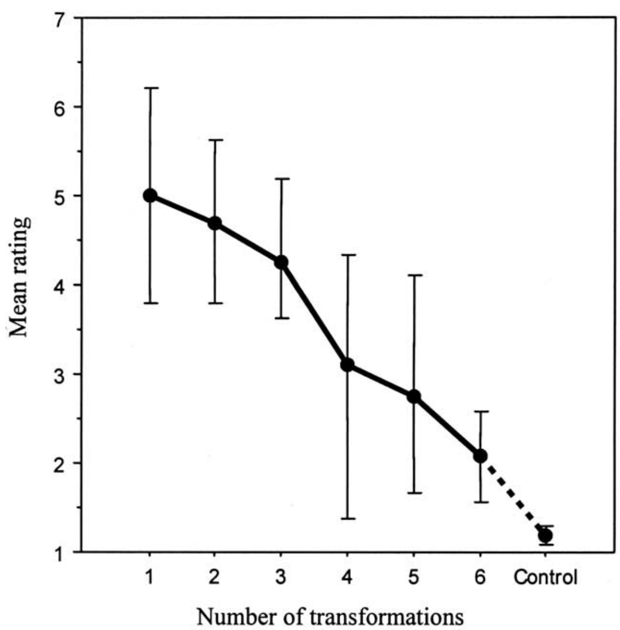


Fig. 7. Experiment 3 (Lego Bricks): Graph depicting the relationship between transformation distance and mean similarity rating, from 1 (very dissimilar) to 7 (very similar). The error bars represent the minimum and maximum ratings assigned by participants.

nature of the materials in Experiment 3 poses a problem for any representation scheme which does not allow structured representations; conversely it lends support to any account such as structural alignment theories to which such structured representations are central. Similarly, the result that number of transformations is a significant predictor of perceived similarity lends support to the general notion of an influence of transformations on similarity, whether or not the particular framework of RD theory is adopted. The support for the role of transformations provided by the results of Experiment 3 moves beyond that of the previous two studies in terms of the basic transformations involved. The initial examination afforded by Experiment 3 suggests that real-world physical constraints influence which transformations are considered. However, further research will be necessary to determine the extent of this influence.

5. General discussion

The results of all three experiments provide first evidence for the claim that similarity should be understood in terms of the general notion of transformation. They meet both evaluation criteria set out for the transformational perspective above: that (1) similarity be shown to be influenced by transformations beyond the restricted set of transformations afforded by featural or spatial models of similarity, thus suggesting the need for a more general theory; and that (2) similarity be shown to be influenced by a wide range of transformations which share nothing apart from the fact that they are all transformations, thus suggesting that the notion of transformation itself is an appropriate building block for this more general theory of similarity.

With regards to the first criterion, the results suggest that classical accounts of similarity are too restrictive. All three experiments provide evidence against purely featural views of similarity. Experiment 1 provides a direct test. The instantiation of the featural model that we tested (i.e. assuming that features correspond to blobs) is, of course, not the most sophisticated featural description possible, given that, in principle, any property including all higher-order regularities such as "symmetry" could be posited as features (Tversky, 1977). Crucially, however, a more sophisticated featural account that succeeds in providing comparable or even superior data fits must not only first be found, it must also be independently motivated. Cognitive science abounds with "mimicry theorems" of all kinds, that is, demonstrations that a particular pattern of data is amenable to explanation by contrasting accounts. It has been argued, for example, that any data indicative of serial processing might be explained by an inherently parallel model and vice versa (Townsend & Ashby, 1983). Much the same has been argued in the debate between imagistic vs. propositional representations (Anderson, 1978), and similar equivalences have recently been highlighted in the context of rule- vs. similarity-based explanations (Hahn & Chater, 1998a).

Given that theories can be stretched beyond all recognition through the addition of suitable post-hoc auxiliary assumptions, a crucial factor in evaluating competing accounts must be not only whether an account can be made compatible with a particular pattern of data but whether it in any way *predicted* it. In this way, Kosslyn (1980, 1981), for example, argued that though propositional representations could be made to fit mental rotation data, there was nothing inherent in a propositional story of mental representation which would have led one to expect lawful effects of rotational distance; all that propositional explanations were doing in this regard was providing post-hoc explanations for a pattern of performance which had logic and coherence only from an imagistic perspective. The same must hold for evaluating featural accounts.

Whether or not *some* featural description that would allow adequate data fits for Experiment 1 is forthcoming or not is almost a moot point in that any such account would seem to necessarily be trying to replicate, post-hoc, in featural terms patterns of behaviour which are predicted only by the more general transformational account. There is nothing whatsoever in featural theories of similarity that would naturally have given rise to the predictions made on the basis of transformations in this experiment.

Ultimately, we suspect that no equivalently performing, independently motivated featural theory could be found. This is due to the basic incompatibility of "features" and "transformations". Though any property of an object can be a feature, transformations are not properties of individual objects. Rather they are *relations between objects*, not – as are features – *inherent properties of objects* themselves. The similarity between a pair of objects is determined by the transformations required to make identical *that pair* of objects. Consequently, the relevant transformations are not inherent in the object itself, but in its relationship with the object against which it is compared. This means that transformations cannot themselves be reduced to features, meaning that featural accounts could achieve the same predictive outcome as a transformation-based account only in some entirely different, unrelated way which just happened to yield the same degree of predicted similarity.

Experiment 2 is just as problematic for a featural (or spatial) account. The sequences of filled circles in Experiment 1 lend themselves naturally to a featural decomposition due to the fact that the "object" is readily parsed into a set of individual circles that are available for featural comparison on a one by one basis. As we outlined above, many of the relevant "features" of the geometric shape stimuli of Experiment 2 become cognitively salient only through contrast between the two comparison objects (for example, the feature "orientation" highlighted by the transformation "rotation"). This contrast, however, is transformationally determined. Arguably, this makes any "features" transformationally motivated. Consequently, transformations again seem explanatorily prior.

The use of simple formations of Lego bricks in Experiment 3 focuses on the central representational weakness of featural accounts – their inability to deal with structured representations and thus adequately to represent relational information. The challenge presented by Experiment 3 is to come up with any even remotely suitable featural description, given the inherent relational nature of materials and transformations.

The limitations of featural accounts suggested by this series of experiments is equally shared by spatial models of similarity, whether they be based on multi-dimensional scaling or standard connectionist networks. From a transformational perspective, featural and spatial accounts of similarity are not wrong; they are simply too restricted. Changes to a feature, feature insertions and deletions as well as changes along a continuous valued dimension are all bona fide transformations; consequently, it is no surprise that both featural and spatial accounts have enjoyed great success in explaining human behaviour. It is merely that the set of cognitively relevant transformations extends beyond this limited set. In summary, our experiments provide evidence that a more general notion of transformation than that embodied in featural or spatial models is required to accurately capture human performance and hence to found a general theory of similarity.

With regards to our second criterion, the wide variety of transformations and their combinations shown to influence similarity in our experiments lends support to the notion that similarity should be conceived of in transformational terms. Specifically, the range of transformations manipulated seem to share nothing except the fact that they are all transformations. This suggests that the notion of transformation might itself provide the right level of abstraction in which to couch a general theory of similarity.

In summary, the experimental results above satisfy both empirical criteria we set out as having to be met in order for similarity plausibly to be thought of as determined by transformational relationships. In this sense, they provide first support for a general transformational approach to similarity. Needless to say, providing evidence sufficient to give a general theory some plausibility is only a small first step. Much else is required before the transformational approach could be viewed as firmly established. In particular, any general theory of similarity must be shown to be successful across the range of possible materials for which lawful similarity judgements can be derived. Though the experiments described in this paper seek to test the theory with a variety of different kinds of stimuli, this selection necessarily constitutes only a small part of the kinds of materials which need consideration and, in this particular respect, they can only provide evidence which is limited accordingly.[8] They do, however, recommend the transformational approach for further study and they suggest that the development of more detailed transformational models of similarity might be worthwhile. There are several considerations that recommend RD as an appropriate starting point for this endeavour of developing a more detailed account. The first is its generality – it can be applied to representations of any kind of information, from visual stimuli, to sentences, to voices, so long as plausible transformations between these representations can be postulated. The second consideration also relates to this generality: many different kinds of specific theory, developed for a particular type of stimulus, may usefully be viewed as special cases of RD – that is, other measures of transformational complexity can be viewed as measures of coding complexity, given an appropriately defined code. Thus, RD represents a broad class of transformational account. Third, RD is based upon a rich mathematical theory of 'information distance' (Bennett et al., 1998), within the field of Kolmogorov complexity theory. Fourth, the RD measure has some attractive psychological properties, such as implying Shepard's well-known 'Universal Law of Generalization' (Shepard, 1987) that relates inter-item confusability to the 'distance' between the representations of those items (Chater & Vitányi, 2002). Finally, though there are many possible alternatives, there are currently no explicit rival theories. Transformational models have been extensively considered in the specific context of shape perception

[8] Regarding future studies, one reviewer felt that "transformations were extremely salient in the present experiments" and that "it is therefore crucial to demonstrate its (transformation distance's) explanatory power for tasks in which transformations are NOT salient". We note that for pairs of items for which no transformational relationship is discernible, RD predicts simply that these items are maximally dissimilar (as would a featural theory for two objects which have no features in common or a spatial theory for items without common dimensions). Item pairs without any clear transformational relationship were included in Experiments 1 and 2 as control items where they did indeed receive the lowest similarity ratings.

and related areas of human and machine vision (e.g. Basri et al., 1998; Burr, 1981; Hildreth, 1983; Hinton, Williams, & Revow, 1992); these models, however, are too specialized to be applicable to the type and variety of stimuli used in our experiments.

With regards to RD theory, the operationalization of transformation distance used in our experiments is a useful first approximation. The number of transformations constitutes only an approximate measure of the Kolmogorov complexity required to turn one representation into another. Crucially, different transformations may have different degrees of complexity – i.e. it is possible that transformations should be ‘weighted’ in some way, rather than implicitly being treated as equal. Should this be correct, then the use of number of transformations will *decrease* predictive accuracy relative to that which would be possible if the relevant complexities were known.

We suspect that the deviations are not substantial for the present experiments. Partly this is because the predictions based on number of transformations turn out to be quite good; partly it is because of the specific design of these studies. Those transformations that we recruit directly from the perception literature are treated as equivalently ‘basic’ or ‘primitive’ in studies of transformations in perception (see Palmer, 1982, 1983, 1991; Shepard, 1970). Moreover, at least in the domain of perceptual inputs that can be represented by one-dimensional symbol sequences, it has been found that, in practice, the complexity measures given by different psychologically-motivated coding schemes typically agree (Simon, 1972). Finally, number of transformations will be a good predictor wherever one is averaging across many specific transformations and their combinations, because on average, two transformations will still be more complex than one, even if individual transformations differ in complexity. For these reasons, number of transformations provides a reasonable proxy for transformational complexity in the present context.

Nevertheless, it would be desirable to be able to make more detailed predictions that incorporate potential differential complexities. We are currently seeking to investigate differential complexities through the examination of perceived similarity for different single transformations. Any resultant information about relative transformational costs derived from this can be used to further refine our predictions and allow more detailed tests.

At present, the ability to make predictions from the RD account of similarity is also limited by our understanding of the psychologically relevant transformations. There is currently no alternative to the strategy chosen here according to which one makes an educated guess as to the relevant transformations in advance of collecting the data. However sensibly chosen that initial guess, it is possible that some of the posited transformations have no psychological validity and/or there are cases where a sequence of transformations can be recast in terms of a single, simple transformation, which is outside the considered set but provides a simple mapping between the representations in question. This latter case will fortunately occur with low probability because the number of multiple transformations is large, and the number of simple transformations into which they might be transformed is small. Thus, by a simple counting argument, the fraction of sequences of transformations that could be recoded simply will be small. While the limits to our current understanding of the psychologically relevant transformations do not prevent transformational accounts from being tested – as is evidenced by our experiments – this will necessarily affect the accuracy of any predictions made, and this will become increasingly relevant as one seeks to differentiate specific transformational models. This makes a better understanding of the psychologically relevant transformations a central topic for future research aimed at detailed transformational accounts, whether these be RD-based or of any other kind. One direction for such future research is to attempt to specify particular sets of transformations that are typically used in particular domains. Specifically, such research should attempt to use more extensively insights from work on transformational theories of perception (Palmer, 1982, 1983, 1991; Shepard, 1981) to tie down the set of transformations appropriate in explaining similarity judgements concerning simple perceptual stimuli. A second important issue in this context is highlighted by the philosopher Goldman. Goldman (1986) pre-figures the notion of similarity as transformation in his work on providing a 'naturalized' (i.e. psychological) epistemology. He raises the empirical question of whether there is a finite set of such transformations or whether the set of cognitively salient transformations is potentially unbounded. The results of Experiment 3 point in the direction of the candidate set of cognitively salient transformations being at least rather large though it is a matter for future research whether other real-world physical manipulations are also salient in this context. Goldman also suggests that the lawfulness of human similarity judgements might be furthered by an inherent *preference ranking* for transformations, which comes into play where multiple transformational sequences could link the same stimulus pair.

Finally, as stated above, a general theory of similarity cannot restrict itself to demonstrations in any single context. Consequently, future research must also seek to apply the transformational approach to different domains. Of particular interest are those areas that appear to require structured representations. As we noted above, RD applies to structured representations in a straightforward way. For example, structured representation of two postures of a hand, in terms of its component palm, figures etc., and a specification of joint angles, etc., can be compared straightforwardly. Given that transformations are likely to be salient in cognitive processing involving motor control, we might expect that 'similar' hand positions will correspond to positions that can readily be transformed into each other (for a transformation-based algorithm applicable to such contour similarities see also Basri et al., 1998). Equally, the linguistic strings *mary sees john* and *john is seen by Mary* may be judged as similar because linguistic operations can turn one into the other (such an account need not necessarily rely on the notion of a transformation, as used in generative linguistics; Chomsky, 1965, 1993). Furthermore, it may be possible to quantify the complexity of transformations between sentences that have just structural, but no lexical, overlap, and so on. A last example concerns understanding the structural similarity between objects or domains that support analogy. An important aspect of the problem of analogical mapping is the problem of finding a connection between two complex domains which exposes the common structure between domains, and hence allows knowledge about one domain to be transformed into knowledge about the other. It may be hoped that RD might provide a measure of the complexity of alternative transformations or alignments between domains.

In summary, considerable research is required before we can hope to possess an empirically adequate, specific transformational account. It is our hope that this paper will encourage such research in the future.

We have presented a new account of similarity, Representational Distortion, according to which the judged similarity between a pair of items is based on the transformation distance between the mental representations of those items. In three experiments, we have provided first evidence for the general claim that the similarity between two objects is determined by transformational relationships. These results present a challenge for other accounts of similarity, based on feature comparison or spatial distance, and they indicate that the view that similarity can be explained in terms of transformation merits further theoretical and empirical investigation.

Acknowledgements

The research reported here was made possible by a grant from the Leverhulme Trust. A preliminary report of these results appeared in the proceedings of the 23rd annual meeting of the Cognitive Science Society. We would like to thank R. Akhtar and A. Weigelt for suggesting configurations of Lego bricks as potential stimulus materials, and Paul Vitányi for theoretical input to this research program. We would also like to thank Todd Bailey, Jacques Mehler and our reviewers for their comments on the manuscript. N.C. was partially supported by European Commission grant RTN-HPRN-CT-1999-00065, the Human Frontiers Science Program, and Oliver, Wyman and Company.

References

Achinstein, P. (1963). Variety and analogy in confirmation theory. *Philosophy of Science*, *30*, 207–227.

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review*, *88*, 249–277.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ashby, F. G., Maddox, W. T., & Lee, W. W. (1994). On the dangers of averaging across subjects when using multi-dimensional scaling or the similarity choice model. *Psychological Science*, *5*, 144–151.

Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: lexical or phonotactic? *Journal of Memory and Language*, *44*, 568–591.

Basri, R., Costa, L., Geiger, D., & Jacobs, D. (1998). Determining the similarity of deformable shapes. *Vision Research*, *38*, 2365–2385.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, *57*, 217–239.

Bennett, C. H., Gács, P., Li, M., Vitányi, P., & Zurek, W. (1998). Information distance. *IEEE Transactions on Information Theory*, *IT-44*, 1407–1423.

Biederman, I. (1985). Human image understanding: recent research and a theory. *Computer Vision, Graphics, and Image Processing*, *32*, 29–73.

Burr, D. (1981). Elastic matching of line drawings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *3*, 708–713.

Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organization. *Psychological Review*, *103*, 566–581.

Chater, N. (1999). The search for simplicity: a fundamental cognitive principle? *Quarterly Journal of Experimental Psychology*, *52A*, 273–302.

Chater, N., & Hahn, U. (1997). Representational distortion, similarity, and the universal law of generalization. *Proceedings of the Interdisciplinary Workshop on Similarity and Categorization, SimCat97* (pp. 31–36). Edinburgh: Department of Artificial Intelligence, University of Edinburgh.

Chater, N., & Vitányi, P. M. B. (2002). Generalizing the universal law of generalization, submitted for publication.

Chater, N., Vitányi, P. M. B., & Stewart, N. (in press). Universal generalization and universal inter-item confusability. *Behavioral and Brain Sciences*.

Chomsky, N. A. (1965). *Aspects of the theory of syntactic structure*. Cambridge, MA: MIT Press.
Chomsky, N. A. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The view from Building 20* (pp. 1–52). Cambridge, MA: MIT Press.
Cooper, L. A. (1976). Demonstration of a mental analog of an external rotation. *Perception and Psychophysics*, *19*, 296–302.
Corballis, M. C., & Roldan, C. E. (1974). On the perception of symmetrical and repeated patterns. *Perception and Psychophysics*, *16*, 136–142.
Dakin, S. C. (1997). The detection of structure in glass patterns: psychophysics and computational models. *Vision Research*, *37*, 2227–2246.
Duncan, J., & Humphreys, G. (1992). Beyond the search surface – visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 578–588.
Feldman, J. (1996). The structure of perceptual categories. *Journal of Mathematical Psychology*, *41*, 145–170.
Fodor, J. A. (1975). *The language of thought*. New York: Cromwell.
Fodor, J. A. (1999). *Concepts: where cognitive science went wrong*. Oxford: Oxford University Press.
Fodor, J. A., & McLaughlin, B. P. (1990). Connectionism and the problem of systematicity: why Smolensky's solution won't work. *Cognition*, *35*, 13–204.
Fodor, J. A., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition*, *28*, 3–71.
Franks, J. J., & Bransford, J. D. (1975). Abstraction of visual patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 65–74.
Garner, W. R., & Clement, D. E. (1963). Goodness of pattern and pattern uncertainty. *Journal of Verbal Learning and Verbal Behavior*, *2*, 446–452.
Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. *Cognitive Science*, *7*, 155–170.
Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199–241). Cambridge: Cambridge University Press.
Gick, M. I., & Holyoak, K. J. (1980). Analogical problem-solving. *Cognitive Psychology*, *12*, 306–355.
Goodman, N. (1972). Problems and projects. Indianapolis: Bobbs Merill.
Goldman, A. I. (1986). *Epistemology and cognition*. Cambridge, MA: Harvard University Press.
Goldstone, R. L. (1994). The role of similarity in categorization: providing a groundwork for cognition. *Cognition*, *52*, 125–157.
Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the non-independence of features in similarity judgments. *Cognitive Psychology*, *23*, 222–262.
Grenander, U. (1996). *Elements of pattern theory*, Baltimore, MA: Johns Hopkins University Press.
Hahn, U., & Chater, N. (1997). Concepts and similarity. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 43–92). Hove: Psychology Press.
Hahn, U., & Chater, N. (1998a). Similarity and rules. Distinct? Exhaustive? Empirically distinguishable? *Cognition*, *65*, 197–230.
Hahn, U., & Chater, N. (1998b). Understanding similarity: a joint project for psychology, case-based reasoning and law. *Artificial Intelligence Review*, *12*, 393–427.
Hahn, U., & Chater, N. (1998c). Real-world categories don't allow uniform feature spaces – not just across categories but within categories too. Commentary. *Behavioral and Brain Sciences*, *21*, 28.
Hahn, U., & Nakisa, R. C. (2000). German inflection: single- or dual-route? *Cognitive Psychology*, *41*, 313–360.
Hahn, U., & Ramscar, M. C. A. (2001). Mere similarity? In U. Hahn & M. C. A. Ramscar (Eds.), *Similarity and categorization* (pp. 257–272). Oxford: Oxford University Press.
Hahn, U. & Ramscar, M. C. A. (2001). Similarity and categorization Oxford: Oxford University Press.
Hampton, J. A. (1995). Testing the prototype theory of concepts. *Journal of Memory and Language*, *34*, 686–708.
Hesse, M. B. (1964). Analogy and confirmation theory. *Philosophy of Science*, *31*, 319–327.
Hildreth, E. (1983). *The measurement of visual motion*. Cambridge, MA: MIT Press.
Hinton, G. E., Williams, C., & Revow, M. (1992). Adaptive elastic models for hand-printed character recognition. *Neural Information Processing Systems*, *4*, 512–519.
Hintzmann, D. L. (1986). Schema abstraction in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.
Hofstadter, D.Fluid Analogies Research Group (1997). *Fluid concepts and creative analogies: computer models of the fundamental mechanisms of thought*, London: Allen Lane.

Hofstadter, D. R., & Mitchell, M. (1994). The COPYCAT project: a model of mental fluidity. In K. J. Holyoak & J. A. Barnden (Eds.), *Analogical connections*. Norwood, NJ: Ablex.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Cognition*, *15*, 332–340.

Imai, S. (1977). Pattern similarity and cognitive transformations. *Acta Psychologica*, *41*, 433–447.

Indurkhya, B. (1989). Modes of analogy. In K. P. Jantke, *International Workshop on Analogical and Inductive Inference 1989* (pp. 217–230). Berlin: Springer-Verlag.

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M. (1981). The medium and the message in mental imagery: a theory. *Psychological Review*, *88*, 44–66.

Leyton, M. (1986). A theory of information structure. 1. General principles. *Journal of Mathematical Psychology*, *30*, 103–160.

Li, M., Li, X., Ma, B., & Vitányi, P. (2001). Normalized information distance and whole mitochondrial genome phylogeny analysis, submitted for publication.

Li, M., & Vitányi, P. (1997). *An introduction to Kolmogorov complexity and its applications*, (2nd ed.). New York: Springer-Verlag.

Luce, P. A. (1986). *Neighborhoods of words in the mental lexicon*. PhD dissertation, Department of Psychology, Indiana University, Bloomington, IN.

Markman, A. B. (2001). Structural alignment, similarity, and the internal structure of category representations. In U. Hahn & M. C. A. Ramscar (Eds.), *Similarity and categorization* (pp. 109–130). Oxford: Oxford University Press.

Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.

McBeath, M. K., & Shepard, R. N. (1989). Apparent motion between shapes differing in location and orientation: a window technique for estimating path curvature. *Perception and Psychophysics*, *46*, 333–337.

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

Mitchell, M. (1993). *Analogy-making as perception: a computer model*. Cambridge, MA: MIT Press.

Mitchell, M., & Hofstadter, D. (1990). The emergence of understanding in a computer model of concepts and analogy-making. *Physica D*, *42*, 322–334.

Mumford, D. (1996). Pattern theory: a unifying perspective. In D. Knill & W. Richards (Eds.), *Perception as Bayesian inference*. Cambridge: Cambridge University Press.

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

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

Oaksford, M. & Chater, N. (1998). Rational models of cognition Oxford: Oxford University Press.

O'Hara, S. (1992). A model of the redescription process in the context of geometric proportional analogy problems. In K. P. Jantke (Ed.), *International Workshop on Analogical and Inductive Inference 1992* (pp. 268–293). Berlin: Springer-Verlag.

Olivers, C. N. L., & Van der Helm, P. A. (1998). Symmetry and selective attention: a dissociation between effortless perception and serial search. *Perception and Psychophysics*, *60*, 1101–1116.

Ortony, A. (1979). Beyond literal similarity. *Psychological Review*, *86*, 161–180.

Osherson, D. N. (1990). Categorization. In D. N. Osherson & E. E. Smith (Eds.), *Thinking: an invitation to cognitive science*. Cambridge, MA: MIT Press.

Palmer, S. E. (1982). Symmetry, transformation, and the structure of perceptual systems. In J. Beck (Ed.), *Organisation and representation in perception* (pp. 95–144). Hillsdale, NJ: Erlbaum.

Palmer, S. E. (1983). The psychology of perceptual organisation: a transformational approach. In J. Beck, B. Hope & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 269–339). New York: Academic Press.

Palmer, S. E. (1991). Goodness, gestalt, groups, and Garner: local symmetry subgroups as a theory of figural goodness. In G. R. Lockhead & J. R. Pomerantz, *The perception of structure: essays in honor of Wendall R. Garner* (pp. 23–39). Washington, DC: American Psychological Association.

Posner, M. I., Goldsmith, R., & Welton Jr., K. E. (1967). Classification of distorted patterns. *Journal of Experimental Psychology*, *73*, 28–38.

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, *83*, 304–308.

Quinlan, J., & Rivest, R. (1989). Inferring decision trees using the minimum description length principle. *Information and Computation*, *80*, 227–248.

Rips, L. J. (1975). Inductive judgements about natural categories. *Journal of Verbal Learning and Verbal Behavior*, *14*, 665–685.

Rissanen, J. (1989). *Stochastic complexity and statistical inquiry*. Singapore: World Scientific.

Ross, B. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology*, *16*, 371–416.

Shepard, R. N. (1970). Second-order isomorphism of internal representations: shapes of states. *Cognitive Psychology*, *1*, 1–17.

Shepard, R. N. (1984). Ecological constraints on internal representation: resonant kinematics of perceiving, imaging, thinking, and dreaming. *Psychological Review*, *91*, 417–447.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*, 1317–1323.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *191*, 701–703.

Shepard, R. N. (1957). Stimulus and response generalization: a stochastic model relating generalization to distance in psychological space. *Psychometrika*, *22*, 325–345.

Shepard, R. N. (1981). Psychophysical complementarity. In M. Kubovy & J. R., Pomerantz (Eds.), *Perceptual Organization* (pp. 279–341). Hillsdale, NJ: Erlbaum.

Simon, H. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review*, *79*, 369–382.

Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*, 3–21.

Townsend, J., & Ashby, F. (1983). *The stochastic modelling of elementary psychological processes*, Cambridge: Cambridge University Press.

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

Ullman, S. (1989). Aligning pictorial descriptions: an approach to object recognition. *Cognition*, *32*, 193–254.

Ullman, S. (1996). *High-level vision: object recognition and visual cognition*. London: MIT Press.

Wallace, C. S., & Boulton, D. M. (1968). An information measure for classification. *Computing Journal*, *11*, 185–195.

Weyl, H. (1952). *Symmetry*. Princeton, NJ: Princeton University Press.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expressions megamix: tests of dimensional and category accounts of emotion recognition. *Cognition*, *63*, 271–313.