Examples And Generalisations: Using Surface Versus Structural Recall Biases To Probe Conceptual Storage

Michael Ramscar, Helen Pain and Sarah Darrington (michael,helen,sarahd@dai.ed.ac.uk)
Department of Artificial Intelligence, University of Edinburgh,
80 South Bridge, Edinburgh EH1, 1HN, Scotland
John Lee (john@cogsci.ed.ac.uk)
Human Communication Research Centre, University of Edinburgh,
2 Buccleuch Place, Edinburgh EH8 9LW, Scotland

Abstract

We argue that the key question in conceptual storage is best viewed not as a question of instances versus generalisations, but rather one of *unitary* versus *multiple* representation accounts of conceptual storage. On previous evidence, it has been difficult to determine whether a particular result stems from stored information regarding the concept or from the processes that operate in invoking a particular concept (Komatsu, 1992). In this paper, we attempt to shed some light on the nature of stored conceptual structure using the different influences that surface and structural features have been shown to have on the recall of a particular representation (Gentner, Ratterman and Forbus, 1993). We conclude that at least some concepts may not be stored using a unitary representation.

Introduction

What is stored in the mind when a person learns a new category? What kind of mental representation (or representations) determine judgements of whether things are to be classed as bicycles, pianos, theories, or games, etc.? These questions lie at the heart of any understanding of human conceptual prowess, and a range of potential answers have been offered in their solution, ranging from necessary and sufficient essences, through defining prototypes, to averages of individual examples: the propositional schematic models of representation used to frame most theories of categorisation have the power to accommodate a combination of defining schemas and/or the examples from which schemas could be constructed with equal coherence, a flexibility that theorists have exploited to the full.

Whilst the work of Wittgenstein (1953) and Rosch, (1978) has tempered enthusiasm amongst researchers for classical essentialist accounts of conceptual storage, there has been much debate about the nature and plausibility of some form of prototype schema as a basis for conceptual storage (see Komatsu, 1992 for a discussion), though what is actually stored is the subject of much debate: Fodor and Lepore (1996) give many good reasons for doubting pure prototype theory, whereby just schemas are stored, whilst Nosofsky (1988) advocates a rejection of stored schemas altogether arguing that only examples are stored, and that new objects are classified by a process of comparison with stored examples. Other researchers favour some kind of theoretical framework as a basis for category storage (e.g. Medin and Ortony, 1989; Keil, 1989)

One point that has eluded many discussions of conceptual storage is that debates about the nature of representations revolve around questions of granularity rather then principle. As Borges (1962) succinctly demonstrates in the much quoted *Funes the Memorius*, even an 'instance' is a generalisation of sorts. When Funes struggles to

> 'comprehend that the generic symbol *dog* embraces so many unlike individuals... it bothered him that the that the dog seen at 3.14 (seen from the side) should have the same name as the dog at 3.15 (seen from the front).' Borges (1962) pp 93-94

he encounters the following problem: if the concept of *dog* determines whether things we encounter are dogs, then in the same fashion the concept *Spot* must similarly unite a certain class of experiences of a particular dog as being experiences of Spot the dog, and so on.

Once this factor is considered, we argue that the key question in conceptual storage is best viewed not as a question of instances versus generalisations, but rather one of *unitary* versus *multiple* representation accounts of conceptual storage. Unitary accounts of categorisation posit a single stored representation - schema, prototype or, perhaps, theory in virtue of which items are classified into a category as the outcome of some process. Multiple representational accounts posit the storage of a number of representations, perhaps at different levels of granularity, from 'instances' (as Malt (1995) argues, regularities - perhaps as a result of perceptual constraints are inevitable at some level) to broad intermediate generalisations, which may jointly or individually result in some object being categorised as the outcome of some process.

Different models of conceptual representation have different implications for theories of categorisation. If a unitary representation model is correct, then one would expect that provided one could specify the stored representation and the process by which objects were related to it, one should in principle be able to give a definitive account of, say why it is some things are X's. On the other hand, a multi-representational account might not admit any definitive account of X's at all, since the multiplicity of relations between the differing stored representations that could influence an object's X-ness might preclude any kind of general account. A specific object's X-ness might be dependent upon that - and only that objects' interaction with a particular subset of the stored elements relating to X-ness, and the concomitant process by which X-ness is adjudged.

A problem in much research into categorisation is that experimental results have rarely - if ever - directly indicated anything about conceptual representation. As Komatsu (1992) concisely notes, it is difficult to determine whether particular results stem from stored information regarding concepts (e.g. propositional or imaginal information) or from the processes that operate in invoking a particular concept (c.f. Smith and Medin, 1981). The vast majority of theories discussed in Komatsu's (1992) comprehensive review assume a 'straight' unitary representation, the exception being granular instance-based approaches to categorisation, and perhaps explanation-based approaches (although the lack of any specific formulation of an explanation (or theory) based model of categorisation makes it impossible to deduce the kind of stored representations such a theory would entail).

In the light of these considerations, we feel that direct evidence regarding the nature of the storage of a concept, or concepts, could have important implications for the way categories are viewed: especially if that evidence fails to support a unitary-representation account.

A process model

In previous work (Ramscar and Pain, 1996) we have questioned an intuitive distinction - held by most psychological researchers - between analogy and metaphor on one hand, and categorisation on the other. We have argued that although one might ordinarily distinguish between category membership and analogy according to realist terms, there are good reasons for treating this distinction with caution at a theoretical level, especially when we focus upon cognitive processes.

The argument for this is twofold. Firstly, in principle: any intuitive, pre-theoretical distinction between real (intra-categorical) and metaphorical or analogical (inter-categorical) cognisance hinges upon a contrast definition that sits ill with what we know empirically and theoretically about categorisation. Analogy and metaphor are consistently defined in contrast to categorisation (e.g. Holyoak and Thagard, 1995); yet a contrast definition relies on an account of the element with which the contrast is to be drawn: the message of many years of empirical and theoretical work (Wittgenstein, 1953; Rosch, 1978; Barsalou, 1983; Murphy and Medin, 1985) is that there is no clear account of intra-categorical associations between individual items with which supposedly inter-categorical associations can be contrasted. Our dubiety towards this distinction is further borne out by analyses of reaction time studies. For instance, Hoffman and Kemper (1987) review of a number of reaction time studies amply demonstrates the absence of evidence for (and amount of evidence against) the widely held belief that literal (intra-categorical) meanings are processed faster than metaphorical (inter-categorical) meanings (Récanati (1995) also reviews some interesting evidence contra the 'two-process' approach).

Further support for this position comes from an earlier study (Ramscar and Pain, 1996) which used Gentner's (1983) Structure Mapping Theory (SMT) to address the question of whether analogy can be distinguished from categorisation by contrasting categorisational with analogical processes.[1] Gentner's theory proposes that in mapping and inference between two representations, systems of relations will be mapped in preference to individual features, and that higher order relations - relations between relations - will be mapped in preference to lower order and first-order (relations between objects) relations: the *systematicity principle* (Gentner, 1983).

Ramscar and Pain's (1996) subjects were presented with Gentner et al's materials and asked to categorise them. It was expected that structure mapping would determine categorisation as well as analogy: i.e. Gentner's theory assumed that match items with only structural similarities ('analogues') belong to different categories: Ramscar and Pain predicted that they would be 'categorised' together. Ramscar and Pain found that a significant majority of the groupings formed by subjects had only shared systematic structure (traditionally defined as analogy) as a common feature amongst members of the categories formed, and that the process underlying subjects' categorising could be modelled - and hence predicted by SMT; they concluded that there was nothing to distinguish subject's analogical from their categorisational behaviour. These results were replicated by Darrington, Lingstadt and Ramscar (1998) who discovered that subjects would even use shared structure in novel object descriptions to *over-ride* existing category groupings.

Our second argument is a pragmatic one. Analogy models can provide a framework through which the interactions between representation and process can be predicted and observed. Traditionally categorisation models have concentrated on object descriptions, making use of very representationally-simple attribute-value lists (Murphy & Medin, 1985), whereas analogy research has examined relationships between highly structured representations (considering the influence of attributes, relations and higher-order relations in judgements of similarity). Research into analogy differs from research into categorisation is in the richness of its process models. A number of detailed, plausible, well accepted models of the analogical process exist (Holyoak and Thagard, 1995; Forbus, Gentner & Law, 1995): the same cannot be said of existing categorisation models. Research in analogy has focused upon the processes of mapping and inference between representations, and the interplay between this process and those representations rather than simply determining what kind of mappings certain static representations allow, a characteristic of much categorisation research (c.f. Komatsu, 1992).

Representation and process

A series of studies by Gentner, Ratterman and Forbus (1993) has explored the factors determining how items are accessed, i.e. how representations are selected in order to allow similarity mapping to take place. These have shown that access relies primarily upon surface attribute (or object attribute) matches, and they propose that the

[1] Though developed as a theory of analogy, SMT is now considered by Gentner to apply more generally as a theory of similarity.

process underlying judgements of similarity can be decomposed into two sub-processes:

- *Accessing* a similar (*base*) situation from memory, based primarily on surface similarity
- Creating a *mapping* from base to target using structural commonalities (SMT).

We believe that this process model can offer a solution to the difficulties, mention above, of determining whether particular effects result from stored representations or from the processes that operate in invoking representations. The following experiment was designed to use this detailed model of the processes that determine the retrieval and mapping (and classifying) of representations to empirically probe the nature of conceptual storage.

The Experiment

In the course of Ramscar and Pain's (1996)[2] classification experiment, subjects were asked to give each of their classes a name that was meaningful to them. Because of the particular nature of that task, this involved subjects developing (and learning) 'categories' that contained some items that had only structural relations in common. By examining the attributes they could recall that were associated with that name, we aimed to use Gentner et al's findings about systematic structures vs. attributes to determine the representation associated with the name. If subjects stored some kind of abstracted prototype i.e. a unitary representation of their category - we would expect that the attributes associated with the most prototypical stories would be most readily retrieved from memory, with attributes recalled insofar as they were relevant to the prototype (perhaps along the frequency lines one might expect from the analysis in table 1).

Story 1 - Base story

Once there was a teacher named Mrs. Jackson who wanted a salary increase. One day, the principal said that he was increasing his own salary by 20 percent. However, he said there was not enough money to give the teachers a salary increase.
When Mrs. Jackson heard this she became so angry that she decided to take revenge. The next day, Mrs. Jackson used gasoline to set fire to the principal's office.
Then she went to a bar and got drunk.

Story 2 - Literal similarity

Professor Rosie McGhee very much wanted a raise. One day the provost announced that he was giving himself a raise. However, he said that since money was short, no one else would get a raise this year.
After Professor McGhee heard this she became so upset that she decided to get even. One hour later, Professor McGhee blew up the administration building with dynamite.

Story 3 - True Analogy

McGhee was a sailor who wanted a few days of vacation on land. One day, the captain announced that he would be taking a vacation in the mountains. However, he said everyone else would have to remain on the ship.
After McGhee heard this he became so upset that he decided to get revenge. Within an hour McGhee blew up the captain's cabin with dynamite.

Figure 1: Sample stories from Gentner, et al (1993) - the text in bold

We did not expect this to happen There is too much empirical and theoretical work that cannot be accommodated by a unitary representation account (Wittgenstein, 1953; Rosch, 1978). Our hypothesis was that subjects would not have abstracted a unitary representation from the stimuli that they had classed together, but would instead have stored a number of representations that they associated with the relevant class name. We predicted that introducing a class name out of context would make it equally likely that any given stored representation associated with that name would be recalled, with the representation recalled initially driving further recall. Since our model of recall is feature driven, we expected results to polarise, with those representations with few surface attributes in common leading to minimal recall of oneanother, and those with many surface attributes in common leading to good recall of oneanother

Subjects

The subjects were 20 volunteers, a mixture of postgraduate and undergraduate students from the Artificial Intelligence Department at the University of Edinburgh.

Materials

Materials were the classified sets of "Karla the hawk" stories (Gentner, Ratterman & Forbus, 1993; see figure 1 for examples) produced by the subjects in Ramscar and Pain (1996). Each set contains 3 stories, and taking one story as a base, the following taxonomy of similarity relationships between the stories in a set can be defined:

- *Literal similarity* (LS) matches to base include both common relational structure and common object descriptions;
- *Structural similarity*, (SS) match based upon a common system of internal relations with base.

Thus each set comprised a base (B), a story literally similar to it (LS), one structurally similar to it (SS - i.e. with no object attributes in common with the base, although SS stories did share some object attributes with LS). In order to determine the relative effects of object versus structural matches in this experiment, the "Karla the Hawk" stories were analysed and rated to determine the level of attribute commonalities between the individual stories in each set.

Two raters gave a numeric value to each of a range of possible surface attribute correspondences between stories (see figure 2 for details), and then individual attribute correspondences were totalled and averaged between the two raters in order to determine the overall correspondences between stories (table 1). Consistency between raters was 82.5%. Differences between raters were resolved by discussion.

Correspondence	Example	Value
1 -1 map	man, man	5 pts
strong map	street, road	4 pts
'analogical' association	conned, robbed fireman, paramedic	3 pts
weak 'analogy'	enlisted, begged fireman, nurse disappeared, shattered	2 pts
weak association	Fred, Bill (both names) fireman, accountant (both jobs)	1 pt

Figure 2: Classification criteria for determining surface similarity.

[2] The experiment discussed here was conducted alongside Ramscar and Pain's study; this is the first time it has been analysed and reported.

Set	B-LS	LS-SS	SS-B
1	100	90	85
2	100	100	85
3	100	65	45
4	100	55	45
5	100	85	55
6	100	85	75
7	100	75	60
8	100	70	45
9	100	85	65
10	100	75	60
11	100	95	75
12	100	75	60
13	100	75	50
14	100	95	70
15	100	50	35
16	100	95	75
17	95	100	70
18	100	80	55
19	100	80	60
20	100	100	60

Table 1: Object attribute (surface feature) similarity ratios between stories by set.

Procedure

During the classification task (Ramscar & Pain, 1996, see above), subjects were asked to give each of the classes they produced a name that 'would be meaningful to them later'. After finishing the classification task, each subject was given a 5 minute break, and then undertook a 20 minute diversionary task (searching for post-codes from a directory) before being given another 5 minute break. The same subjects were then presented with a sub-set (usually 4) of the names they had assigned to classes during the classification task, and asked to 'write down what you can remember about the various features (or you may like to see them as attributes) of each of the scenarios associated with each name. E.g. you may have had a scenario about a door that needed varnishing. Features, or attributes, associated with such a scenario would be "door" and "varnishing"'. Subjects were given 10 minutes to complete the task.

Results

The 20 subjects yielded a total of 70 recall episodes.

Scoring

The recalled features were evaluated by two judges using the same scale that was used to evaluate feature correspondences between stories (see figure 2). Points were awarded for correspondences between features named by the subjects and the actual feature name in each given story in order to get a reflection of the accuracy of each individual recall episode. The total attribute recall for each story was calculated and averaged between raters. Consistency between raters was 84.1%; as in the rating of story commonalities, differences between the raters were resolved by discussion.

Individual Story Recall

B was best recalled story for 37% of all recall episodes, SS in 30% of cases and LS in 33%; as predicted, there was no significant bias towards recalling any particular type of story. However, when we looked at the pattern of recollection, irrespective of the particular stories each subject had recalled, there were significances in the quality of recall between the best recalled and the next best, and the next-best and the worst recalled stories. Subjects tended to clearly recall one story better (70 cases, Mean (M) = 19.75) than the next (M = 14.74), (within groups $t(69) = 8.846$, $p < 0.0001$), and then these next best recalled stories better than the worst (M = 11.1), $t(69) = 11.802$ $p < 0.0001$.

Individual Recall Orderings By Story Type

Base (B)

In cases where B was the most recalled story in terms of features (26 cases, M = 18.44), there was a significant difference in the quality of recall over the next best recalled story type, LS (M = 14.83), $t(25) = 4.434$ $p < 0.0001$, which in turn was recalled significantly more than SS (M = 10.15), $t(25) = 3.77$ $p < 0.0001$.

Literal Similarity (LS)

When LS was the best recalled story (M = 20.78), next best was B (23 cases, M = 15.85), $t(22) = 5.288$ $p < 0.0001$, with SS recalled least well (M = 11.96), $t(22) = 2.095$ $p < 0.05$.

Structural Similarity (SS)

In cases where SS stories were qualitatively best recalled (21 cases, M = 20.12), the next best recalled type was LS, (M = 13.43), $t(20) = 5.886$ p = < 0.0001, and the least recalled type was B (M = 11.36), although the between type difference between LS and B was not significant ($t(20) = 1.901$ $p < 0.072$); the difference in recall quality between SS and B was still significant ($t(20) = 6.609$ $p < 0.0001$).

B Best Recalled	Next Best - LS	Next Best - SS
26 cases; Mean = 18.44	Mean = 14.83	Mean = 10.15

LS Best Recalled	Next Best - B	Next Best - SS
23 cases; Mean = 20.78	Mean = 15.85	Mean = 11.96

SS Best Recalled	Next Best - LS	Next Best - B
21 cases; Mean = 20.12	Mean = 13.43	Mean = 11.36

Table 2: Mean recall orderings by story type

Discussion

The experiment produced little evidence to support a belief that our subjects had abstracted and stored schemas from the groups they had classified, despite the fact that a shared structural schema was the basis of subjects' original classification decisions (Ramscar & Pain, 1996). If some version of a stored prototype theory were true, we expected a majority of LS features to be recalled in most instances. In fact, B features were most often recalled, though not significantly: the trend favoured a random distribution. Another result that might also favour prototype theory would have been a situation where all the stories were recalled with much the same frequency, i.e. LS=B=SS, since such a result could be a product of the strong feature commonalities between the LS stories and members of both of the other story types. However, there was a significant trend for subjects to recall one story more than another, and the next best story more than the least recalled story.

One criticism of this initial conclusion might be that our task, by asking subjects to write down features associated with individual scenarios, biased against recall of a single unitary or abstracted representation. However, again, given that LS story set attributes formed the intersection of the sets of features for *all* of the story sets, we would expect any unitary or abstracted representation of all of the stories in a given set to contain mainly LS story set attributes and to cause a majority of LS story set attributes to be recalled. Clearly, this was not the case.

If subjects randomly recalled an individual instance of a class then we would expect from Gentner et al's similarity recall findings that the attributes of this instance should influence which other story they might recall from the class: if B is recalled, recall of a B story should prompt recall of an LS story over a SS story, as B shares more surface attributes with LS than SS (B and LS share 10 surface attributes to every 6 shared by B and SS), and Gentner et al's findings were that surface attributes rather than shared structure promote recall: where B features were best recalled this pattern emerged throughout our study. As we predicted, as a result of LS stories sharing a higher percentage of surface attributes with B than SS shared with B, recall of B led to a significantly higher quality of recall for LS than SS. The results where LS stories were most strongly recalled also supported this analysis, with SS recall prompting significantly better recall of B attributes than SS attributes (even though, comparatively, LS shared more surface features with SS than SS-B).

Indicative of the fact that SS shared fewer surface commonalities with the other story types, (SS-LS attribute commonalities were much weaker than B-LS, and in specific sets little greater than B-SS; see table 1), good SS recall did not produce a bias towards LS or B as the next-best recalled story type; although our results showed some tendency towards SS prompting LS over B, it was not significant (see also table 2).

Our hypothesis that subjects would recall stories individually from memory is further supported by the nature of subject's recollection. Irrespective of the particular stories each subject had recalled, there were significances in the quality of recall between the best recalled, and the next best, and the next-best and the worst recalled stories.

A possibility that we cannot eliminate in this instance is that subjects may have been influenced by the order in which stories were presented. In our study, presentation was randomised in a manner that made it impossible to correlate presentation and recall orderings. However, as we noted above, there is a very strong correlation between the features of the LS story in each set and the features of an idealised prototype of that set (i.e. LS features, especially at the level of names of relations, are most typical of all members). Since subjects, who used shared structure to group the stories, did not show a tendency to recall LS features best, any correlation between presentation order and recall order could only further support the case against some shared representation having been abstracted and stored.

Applying Gentner, et al's (1993) persuasive analysis of the influence of surface vs. deep structure on recall to our results, it would appear that being presented with a class name in no particular context caused the subjects to randomly recall one of the examples associated with that name, and then use that example as the stimulus for recalling other class members. On this evidence, it would appear that subjects had stored class examples along with a cue - the class name rather than any generalisation of the class itself. These findings suggest that at least some concepts are stored as multiple-representations, rather than as unitary conceptual schemas.

General Discussion

We have shown in this paper that subjects if subjects name a class whose membership is determined entirely by shared structure, they can retrieve information regarding its members without appearing to abstract a common schema definitive of that class. Such evidence for a non-unitary account of category representation should not come as a surprise: the literature is filled with material that casts doubt of the plausibility of unitary accounts of concept representation (Wittgenstein, 1953; Rosch, 1978; Komatsu, 1992).

Our interpretation of these findings does not however lead us automatically to the view that 'category' associations are exclusively driven by examples; rather, we subscribe to the view put forward by Wittgenstein (1953; see also Ramscar, 1997). Wittgenstein provides a number of good reasons to believe that human 'categories' cannot be given simple unitary accounts that are amenable to definition by a single schema; he also questions what could be intrinsic to a 'generalised' category schema that would cause it to be used differently to an example of that which it was supposed to be a generalisation of. We find these questions compelling. At the same time, experience, intuition and a respect for natural parsimony makes us wary - though in practice rather than in principle - of jumping to the conclusion that all categorical associations are a sum of processed similarities with all stored exemplars. We would argue, however, that progress is not contingent on finding a solution to the question of exactly how categories are represented. It is the *process* of association that is important to us, and from a processing point of view, the question of whether a stored schema relates to an example or some intermediate generalisation is not necessarily important. What *is* important, from the point of view of processing, is to show how stored representations (whatever their exact nature) relating to the individual surface elements (the feature 'nodes' in SMT) comprising the schemas that represent items being associated in the manner described in structure mapping theory are recursively processed in turn, such that the 'network of similarities that determines a given association can settle (see Ramscar, Pain and Cooper (1997) for an illustration of our recursive view of the associative mapping process).

Another advantage of viewing categorisation from a process oriented perspective is that it offers up the possibility of bridging the divide between two seemingly conflicting accounts of categorisation: similarity-based accounts, and more rationalist 'theory-based' accounts. Whilst similarity-based accounts of categorisation (e.g. Rosch, 1978) can capture much of the nature of our long-term categories (that they have no explicit definitions, and that there are usually a number of properties that are generally associated with 'categories'), they have been criticised for failing to capture the explanatory power of categories, and for failing to explain why people may categorise in ways that go beyond surface similarity (Medin and Ortony, 1989). Keil (1989) notes that whilst children's concepts are based upon pure similarity 'original sim' - these then get replaced with more theoretically based conceptual understandings as a child develops (see also Murphy and Medin, 1985; Medin and Ortony, 1989).

What these theory based accounts lack is any clear description of the process by which theories are supposed to govern categorisation, or, importantly, exactly what kind of theoretical understandings are supposed to underpin human conceptual understandings which are notoriously vague and often incoherent. We believe that Gentner's (1983) insight, that relations between features (i.e. *constraints*) can drive computations of similarity, can go at least some way towards squaring this apparent circle. The assumption that similarity cannot account for the apparent theoretical notion of categorical associations stems from the assumption that similarity is driven by features alone. When the relational constraints that drive the similarity comparison process are considered, and once a richer notion of mental representation than unrelated feature clusters is included in the picture, it is possible to see how similarity *can* provide an account of the associative process that determines human categorisation decisions which capture their theoretical flavour,[3] *without* entailing theoretical accounts of individual 'categories' which by their nature (c.f. Wittgenstein, 1953) are more of the nature of underdetermined collections of generalisation, similarity and history than they are unitary 'theoretical' entities.

Acknowledgements

The authors would like to thank: Richard Cooper and two anonymous reviewers for their comments; and Dedre Gentner for her assistance in providing our materials.

References

Barsalou, LW (1983) Ad hoc categories *Memory and Cognition*, **11**:211-227

Borges, JL (1962) *Labyrinths* Penguin, New York, NY

Darrington, S, Lingstadt, T and Ramscar, MJA (1998) Analogy as a sub-process of categorisation. *Submitted.*

Fodor, J & Lepore, E (1996) The red herring and the pet fish: why concepts still can't be prototypes. *Cognition* **58**:253-270

Forbus K, Gentner, D, & Law, K (1995) MAC/FAC: A model of similarity based retrieval *Cognitive Science* **19**:141-205

French, RM (1995) *The Subtlety of Sameness*, MIT Press, Cambridge, Ma.

Gentner, D (1983) Structure-mapping: a theoretical framework for analogy. *Cognitive Science*, **7**: 155-170.

Gentner, D, Ratterman, MJ & Forbus, K (1993) The roles of similarity in transfer. *Cognitive Psychology* **25**: 524-575

Hoffman, R R and Kemper, S (1987) What could reaction-time studies be telling us about metaphor comprehension? *Metaphor and Symbolic Activity* **2(3)**, 149 - 186

Holyoak, KJ and Thagard, P (1995) *Mental Leaps*. MIT Press, Cambridge, Ma.

Keil, FC (1989) *Concepts, kinds and cognitive development.* MIT Press, Cambridge, Mass.

Komatsu, LK (1992) Recent views of conceptual structure. *Psychological Bulletin*, **112(3)**, 500-526

Lakoff, G and Johnson, M (1980) *Metaphors we live by*. University of Chicago Press, Chicago, Ill.

Malt, BC & Smith, EE (1984) Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behaviour* **23**:250-269

Malt, BC (1995) Category coherence in cross-cultural perspective. *Cognitive Psychology* **29**:85-148

Murphy, GL and Medin, DL (1985) The role of theories in conceptual coherence. *Psychological Review*, **92**, 289 - 316

Nofosky, R M. (1988) Similarity, frequency and category representations. *Journal of Experimental Psychology: Learning Memory and Cognition* **14**:54-65

Ramscar, MJA & Pain, HG (1996) Can a real distinction be made between cognitive theories of analogy and categorisation? (*In Proc.*) *18th Annual Conference of the Cognitive Science Society.* LEA, New Jersey.

Ramscar, MJA, Pain, HG & Cooper, R (1997) Is there a place for semantic similarity in the analogical mapping process? (*In Proc.*) *19th Annual Conference of the Cognitive Science Society.* LEA, New Jersey.

Ramscar, MJA (1997) Wittgenstein and the nature of psychological categories *Proceedings of SimCat 97*, Dept. of Artificial Intelligence Conference Proceedings, University of Edinburgh, Scotland, 205-211.

Récanati, F (1995) The alleged priority of literal interpretation *Cognitive Science* **19**:207-273

Rosch, E, Simpson, C. & Miller, RS (1976) Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance* **2:4**:491-562.

Rosch, E (1978) *Cognition and Categorisation.* Lawrence Earlbaum Assoc., New Jersey.

Smith, EE, Shoben, EJ, & Rips, LJ (1974) Structure and process in semantic memory: a feature model for semantic decisions *Psychological Review* **81**:214-241

Smith EE and Medin D (1981) *Categories and concepts.* Harvard University Press. Cambridge, Mass.

Wittgenstein, L *translated by* Anscombe, E (1953). *Philosophical Investigations* Blackwell, Oxford.

[3] Keil (1989) rejects SMT as a basis for such an explanation because it gives no privileged status to the *causal* constraints that he argues are most important to category judgements. We find this mystifying, since SMT maps according to the most weighted constraints embodied in representations of items in working memory: if causal constraints are privileged, one would expect this to be reflected in the weighting of mental representations. It is a strength of SMT that by *not* favouring causal constraints it does not preclude categorical associations formed on the basis of other constraint systems.