*Psychologica Belgica*
1995, 35-2/3, 81-84.

# SIMILARITY AND CATEGORIZATION: INTRODUCTION

André VANDIERENDONCK
*University of Ghent*

Jean-Pierre THIBAUT
*University of Liège*

In this introduction the contributions of the present thematic issue are situated with respect to four different aspects of the similarity notion: the computational basis of similarity, constraints on similarity judgments by processes of selection and context-sensitivity, flexibility versus stability of similarity and the relationship of similarity to general background knowledge.

For many people there is no big difference between a judgment of the similarity of a number of things and categorization of the objects into a number of groups. In both tasks similarity among the objects is the basis for the choices made.

Many present-day theories of concept representation take this relationship between similarity and categorization for granted. It is difficult to think of an explanation of categorization behavior in which the similarity notion would be absent. Even the older rule-theories or hypothesis-testing theories (Bruner, Goodnow & Austin, 1956; Hunt, 1962) rely implicitly on similarity: things that share relevant attributes belong to the same category and may be considered to be similar in this respect.

In the seventies, the similarity notion has been given more weight, and since then, theories of category learning incorporate inter-object similarity in one way or another (e.g., Barsalou, 1983; Kruschke, 1992; Medin & Schaffer, 1978; Nosofsky, 1984; Posner & Keele, 1968; Rosch, 1975a, 1975b). The present issue of *Psychologica Belgica* elaborates the relationship between similarity and categorization, twenty years after the publication of Rosch's seminal work on a prototype structure in the representation of natural kinds (Rosch, 1975a, 1975b). This past twenty years has witnessed quite some changes in the way concepts have been viewed and in the techniques used to study categorization and concept representation.

The collection of papers presented here, all try to clarify one or more aspects of the relationship between categorization and similarity, and they do this in different ways. It is obvious that the problem can be addressed by looking more closely at similarity or by looking more closely at categorization or even by directly analyzing the interaction of similarity and categorization.

---

Correspondence concerning this article should be addressed either to André Vandierendonck, Department of Experimental Psychology, University of Ghent, H. Dunantlaan 2, 9000 Ghent (electronic mail may be sent to Andre.Vandierendonck@rug.ac.be) or to Jean-Pierre Thibaut, Laboratory of Psycholinguistics, University of Liège, Department of Psychology, 5, Bd du Rectorat, 4000 Liège (electronic mail may be sent to jthibaut@vm1.ulg.ac.be).

It will be clear for the reader that the authors in the present volume differ in their definitions of similarity. All authors appear to accept that similarity is computed on the basis of more primitive information. This *computational aspect* is the main theme in the paper by Iven Van Mechelen and Gert Storms, *Analysis of similarity data and Tversky's contrast model*. They see similarity as the result of a process in which information from different sources is integrated and they discuss the validity of techniques that reconstruct a (geometrical) representation of the set of objects. Their analysis centers on the utility of a number of data-analytic methods as a means to answer the criticisms formulated by Tversky (1977) on some similarity measures in the context of his contrast model.

The article by James Hampton, *Similarity-based categorization: The development of prototype theory*, also pays ample attention to the computational aspects of similarity. Not the measurement of similarity but its explanatory power in the context of prototype theory is the focus of his contribution. He addresses the question which constraints are necessary for similarity computations to explain some common observations, such as the findings that concepts do not have a clear conjunctive definition, that concepts are associated also with non-defining information, and that concept hierarchies are sometimes nontransitive.

While all the authors in the present issue agree on the computational nature of the similarity concept, some authors have chosen to focus on the elements that enter the computation. This interest, also present in Hampton's article, is clearly evident in the paper by Jean-Pierre Thibaut and Philippe Schyns, *The development of feature spaces for similarity and categorization*. They propose that similarity depends on a selection and interpretation of information required to identify features and object properties in the first place. After identification of the features, apparently, similarity is the product of a computational process. Likewise, Gregory Murphy and Thomas Spalding, *Knowledge, similarity and concept formation*, stress the importance of selection in the process of similarity computation. Where Thibaut and Schyns are in search of constraints at the level of knowledge about objects and their features, Murphy and Spalding call upon the influence of more general knowledge and theories about the domain.

The notion of similarity can be one of stability or one of flexibility. It is well known to researchers of similarity that the judgments provided by the participants in an experiment are not absolute. Small changes in the procedure or the instructions can yield rather different similarity estimates. To some theorists this is a fundamental characteristic of similarities; to others the issue is merely methodological. Robert Goldstone, *Mainstream and avant-garde similarity*, believes in the flexibility of similarities. He reviews empirical evidence that demonstrates flexibility and context-sensitivity of similarity judgments. In his view, however, this flexibility is restricted, and he goes on to review some

conditions under which stability prevails.

Although the issue of flexibility is the main theme in Goldstone's contribution, he is not the only one to view similarity as flexible. In fact, Thibaut and Schyns also stress the flexibility of similarity, but they see this as a product of the task content and of the history of categorizations a subject has made. Murphy and Spalding's position is similar, but again similarity perception is the product of the influence from theories about the domain.

Since Murphy and Medin (1985) so eloquently defended a position in which theories about the world, implicit cognitive models (see also Lakoff, 1987), and the like, affect categorization and similarity perception needed for categorization, more and more studies have concentrated on the influence of general background knowledge and of specific knowledge on similarity perception. Much of this line of investigation can be found in the article by Murphy and Spalding, but it is also manifestly present in the paper by Goldstone. Presently, it is a matter of debate whether the kind of attribute knowledge stressed by Thibaut and Schyns falls within the range of the general category intended by Murphy and Medin. One possible problem that will have to be solved in this respect, is the status of features. Sometimes features are given as perceptual or as visuo-spatial information; at other occasions features are derived from perceptual information and may be concepts that are used as a reference for deciding category membership of objects. When features are concepts, they belong to the knowledge store. However, to possess a theory about a domain also implies that the knowledge base contains structures or schemata or models that relate one feature concept to an-other one. It is clear, then, that as research moves on, a more precise formulation of the notion of background knowledge and implicit theories will be required to clarify the similarities and the differences between the positions taken by Thibaut and Schyns on the one hand and by Murphy and his colleagues on the other hand.

Due to their focus on data-analytic methods, Van Mechelen and Storms do not express a clear statement concerning this question, although the final part of their article shows that they are well aware of these problems. The issue is also beyond the scope of Hampton's paper, but in the concluding part of his paper, he makes his position more explicit:

> Concepts may be deeply embedded in theory (Murphy & Medin, 1985), but there may be some advantage in retaining a simple representational form for individual concepts, and placing more complex information in a higher level general knowledge memory store. (p. 123)

This is probably a position that—implicitly—has been taken by many researchers of concept learning, and which is probably still consistent with Murphy and Medin's position.

In summary it may be said that the papers collected in this thematic issue discuss the problems of the relationship between similarity and categorization

from a number of different angles, while giving attention to a number of theoretical issues: the measurement of similarity, the basic flexibility of similarity, the importance of selection processes in similarity judgment, the impact of object and feature knowledge, the influence of general background knowledge, etc. In other words, this issue gives a view on some today's questions in categorization.

## References

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

Bruner, J.S., Goodnow, J.J., & Austin, G.A. (1956). *A study of thinking*. New York: Science Editions.

Hunt, E.B. (1962). *Concept learning. An information processing problem*. New York: Wiley.

Kruschke, J.K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*, 22-44.

Lakoff, G. (1987). Cognitive models and prototype theory. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 63-100). Cambridge: Cambridge University Press.

Medin, D.L., & Schaffer, M.M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207-238.

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

Nosofsky, R. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 104-114.

Posner, M.I., & Keele, S.W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*, 353-363.

Rosch, E.H. (1975a). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192-233.

Rosch, E.H. (1975b). The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 303-322.

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