Real-world categories don't allow uniform feature spaces – not just across categories but within categories also

Ulrike Hahn and Nick Chater

Department of Psychology, University of Warwick, Coventry CV7 4AL, United Kingdom. **u.hahn@warwick.ac.uk**
n.chater@warwick.ac.uk

Abstract: The Schyns et al. target article demonstrates that different classifications entail different representations, implying "flexible space learning." We argue that flexibility is required even at the within-category level.

We welcome this timely emphasis on the need for "flexible spaces" in categorization. In this commentary, we ask how far the points made must be taken. The target article stresses that new features are required by new categorizations and that assuming a single, fixed, object representation suited for all possible classifications is unrealistic. Our work on similarity-based categorization (Hahn 1996; Hahn & Chater, in press) has stressed that fixed, uniform representations are inappropriate even *within* a category. Where Schyns et al. emphasize that, for example, a single fixed-length vector for object representation is insufficient across categories, our aim has been to show how a single fixed-length vector representation is overly restrictive even for a single category. That "uniform feature spaces" are insufficient even within a category becomes apparent with the analysis of real-world materials such as legal cases (Hahn 1996), but more real-world categories suffice equally: imagine encountering a particular chair, one with a back rest and four legs; the next exemplar encountered might have armrests too – a new "dimension" that comes into play only at this point, yet another chair might have a swivel base instead of four legs, and so on. The "feature space" for the category emerges only gradually as more and more examples are encountered. The crucial point, however, is that for many categories, if not most, it is never definitively fixed. New, previously unanticipated variations can arise all the time. The problem is not simply that of encountering a sufficient number of exemplars to allow determination of the space of possibilities, because this space generally is not bounded (at least from the agent perspective). This follows from considering a key difficulty for rule-based systems, that rules – whether attempting to govern everyday, commonsense knowledge or specialist domains such as law – almost always admit of exceptions (Hahn & Vogel 1997; Oaksford & Chater 1991; Reiter 1980). These exceptions, which are both unforeseeable and too numerous to allow enumeration in advance, require the ability to perform nonmonotonic or default reasoning in rule-based contexts. But that potentially relevant features are not exhaustively known in advance does not just affect rules and rules alone. They are equally unavailable for any mode of organizing conceptual knowledge. Thus, realistic models of categorization must allow representation and evaluation of "novel" features.

That this is not just a pedantic point that can be ignored in practice is documented by work in machine learning and artificial intelligence. The problem is well known in the context of rule-based systems (Reiter 1980), but instance-based approaches to classification in machine learning have also recognized the need to confront the problem of "novel attributes" (Aha 1992). The aim of this research is to build classification systems that work with practical problems, not ambitious cognitive models. Cognitive modeling should treat the issue all the more seriously.

There is a serious problem, then, for any account of categorization that assumes fixed representations, whether this strait-jacket of uniform representation stems from practical considerations about representation and learning procedures (e.g., backpropagation networks) or stems from the very nature of theory (e.g., spatial models of similarity).

Our own approach to similarity and categorization is based on the notion of transformation between objects, a general concept that encompasses similarity as "feature-overlap" or as distance in similarity-space as a special, restrictive case (Chater & Hahn 1997; Hahn & Chater 1997). Similarity between objects is assumed to depend on the ease of transformation of the representation of one object into representations of the other. Psychology has seen transformational accounts of similarity advanced in the past (Franks & Bransford 1971; Imai 1977). Our account of "representational distortion" provides a foundation in terms of the notion of Information Distance from the branch of algorithmic complexity theory known as Kolmogorov complexity (Li & Vitanyi 1993).

Crucial for the present context is the concept that similarity assessment no longer conceives of objects as residing in a feature space, but instead in transformation space. Features are only of interest as the objects of transformations; in this sense, the account is independent of particular features. As a consequence, there is no need for the same set of features to be present throughout. Also, the same features can be the object of different transformations as these arise from the particular pair of stimuli under consideration. The search for transformations itself influences the features found; consequently, the same basic features can give rise to different stimulus descriptions as a function of the particular comparison.