New-feature learning : How common is it?

(French, R. M. and Weaver, M. (1998). New-feature learning: How common is it? (Commentary) *The Behavior and Brain Sciences 21*(1), NJ: LEA, 26)

Robert M. French
Psychology Department
University of Liège, Liège, Belgium
rfrench@ulg.ac.be

Mark Weaver
Corvus Development, Inc.
Ann Arbor, MI
weaver@corvusdev.com

Abstract

The fixed-feature viewpoint the authors are opposing not a widely held, theoretical position, but rather a working assumption of cognitive psychologists — and thus a straw man. We accept their demonstration of new-feature acquisition, but question its ubiquity in category learning. We suggest that new-feature learning (at least in adults) is rarer and more difficult than the authors suggest.

The authors are, in the main, justified in their skepticism about conventional approaches to category learning—to work with conveniently small numbers of fixed, high-level features is to make a set of simplifying assumptions that have the potential of missing much of the complexity of the category learning process. However, we find their supposed opposition to be something of a straw-man and their conclusions go beyond the supporting the evidence. Despite this, the work they present has important implications which the authors themselves have only begun to explore and provides valuable insights into the relationship between the learning of new categories through the acquisition of the feature sets.

To begin where the target paper does, the authors contrast their position with a fixed-feature alternative which they describe as “an influential approach”. The problem with this is that they are not clear about the difference between fixed features as a theoretical position versus fixed features as a working assumption. As a theoretical position, a “fixed-feature stance” is clearly an untenable straw-man. Features must, of course, come from somewhere and since it is surpassingly unlikely that we are born with all the features we’ll ever need (like the eggs in baby girl’s ovaries), one must presume that features are, in fact, learned. However, there is no evidence that the “fixed-feature stance” has been seriously put forward as a theoretical position. It is, rather, a (widespread) working assumption. Is this assumption valid? The authors argue that it is not. However, even if we accept the authors’ claim that they have demonstrated the acquisition of new features during category learning — which we do — *it does not follow that the acquisition of new features is a* ***typical*** *part of category learning.* In fact, it is reasonable to expect that, just as adult speakers of English have in place all the phonemes they will ever need to represent English words (even the many thousands outside of their current vocabularies), they also have in place all the primitive features needed for learning new categories in many domains. What is more, the exceptional character of the stimuli in the authors’ "Martian Cells" and "Martian Rocks" experiments is a tacit acknowledgment of how hard it is to force subjects outside their existing areas of "featural competence".

Thus, we argue that the widespread assumption of fixed features in category learning experiments *is* supported in the general case of adult category learning.

One of the most promising but elusive points in the target article involves the distinction between "primitive" and high-level, "functional" features. On the one hand, the authors point out that almost *any* feature is composed of lower-level features—at least until one gets all the way down, for example, to the edge detectors in low-level visual cortex. This makes the distinction between "primitive" and "functional" features difficult, if not impossible, to maintain. On the other hand, the authors hang much of their argument on the distinction:

> "If a primitive set of features can capture all the regularities and categorizations accommodated by the functional features…then the hypothesis that feature creation is needed to allow a system to represent object properties it was previously incapable of cannot be maintained." [p. 10]

One way out of this apparent contradiction to think of both "primitive" and "structured" features as hierarchically organized building blocks from which new categories and concepts may be constructed. As we've argued, since many categories will be able to share the same general-purpose building blocks (e.g. something like Biederman's (1987) geons), category learning can often proceed by assembling existing building blocks. In other cases (as with the "Martian rocks" and "Martian cells", we suspect), new building blocks are needed, but they are few in number and, so, may be acquired as a side effect of category learning without significantly impeding the process. At the other extreme—something the authors have not acknowledged—are situations where appropriate features are generally missing. In addition to the case of infants confronting a bewildering world, one might also add those cases where adult learners attempt to master new and unfamiliar domains of knowledge. Examples of this kind of feature learning might include the experience of a native English speaker learning to hear the tonal phonemes of Mandarin Chinese or of a neophyte mushroom gatherer learning to discriminate edible and poisonous varieties). Feature (and, therefore, category) learning in such cases is often slow, difficult, and even painful.

Finally, we suggest that the authors may have, ironically, focused their demonstration of new-feature learning precisely where it most difficult to observe—namely with readily perceived, concrete objects. Both evolutionary constraints (favoring concrete categories) and universally shared experience with objects in the world suggest that by the time of adulthood, most of the primitives humans need to handle concrete categories will already have been learned—which is why we argue that a fixed-feature approach for the learning of concrete categories is a reasonable working assumption. On the other hand, as categories grow more abstract (as in chess playing, mathematics, classical music, architecture, aeronautics, etc.) – and therefore less universal and directly related to needs of survival in the physical world – one might expect a far greater degree of new-feature creation to be observed. Thus, the slow building of expertise in abstract domains may present a better opportunity for studying the interaction between feature acquisition and category learning.

In conclusion, we feel that the authors have, needlessly, presented their argument in contrast to a straw-man opponent. Although, the widespread working assumption of fixed-features in category learning *is* justified, we agree that the acquisition of new

features is an important topic which has to date received scant attention in both empirical research and category learning theory. The target article represents a useful start in both of these areas.

References

Biederman, I. (1987) Recognition-by-components: A theory of human image understanding. *Psychological Review, 94* (2), 115-147

Acknowledgments

This work was supported in part by Belgian PAI grant No. P4/19.