



## SHAPE BIAS SPECIAL SECTION COMMENTARIES

### The shape of things to come: the future of the shape bias controversy

Frank C. Keil

*Department of Psychology, Yale University, USA*

#### Introduction

When confronted with a controversy such as the reason for the shape bias, one wonders, as Rodney King once did, ‘Why can’t we all just get along?’ Is there some combination of theory and data that could be acceptable to all parties? Often such conflict resolutions come from showing the parties are talking at cross purposes and that, once one clarifies terms and meanings, they can all live together in one big happy theoretical family. In reading over these papers, I kept trying to find such a resolution, trying to see if there was a way to talk about what is going on that everyone would agree with, with the important side benefit of not alienating half the authors. In the end, I’m not sure such a happy ending is possible. There seem to be real disagreements at work here that cannot easily be defined away. Those disagreements, however, may have less to do with the shape bias in particular and more do with very general approaches to learning and cognitive development. It is important to understand where the key disagreements lie, because finding their real locus is central to knowing why a broader community beyond the ‘shape bias people’ should care about the controversy.

#### Why should we care?

At first glance, it may not seem to matter much who is right in this controversy. What difference does it make if the shape bias reflects associations between word labels and object shapes as opposed to beliefs about their kinds of properties? Either way, it is clear that young children will often rely on shape in trying to figure out how to extend word meanings. Perhaps the controversy does not really matter. To better see what is at stake, it is useful to try to find the common ground that all parties to the disagreement share. We can then understand what really matters to the various parties. To differing degrees all of these papers make some effort at establishing this common ground, but it is still important to try to do this with a bird’s eye view of all of them at the same time. Here are the top ten points of agreement.

1. Young children will often extend word labels in ways that reflect the shape of the original taught-on object.
2. Young children do not invariably extend labels on the basis of shape and will sometimes use other dimensions and attributes of objects.
3. There are developmental changes in naming patterns and in how those patterns are related to the shape bias.
4. Local contexts can influence behavior. In particular they can influence labeling, and the degree to which the shape bias is manifest.
5. Associative learning is real.
6. There are non-associative components to human thought as well.
7. Bounded objects are where the shape bias is most evident.
8. Microgenetic longitudinal studies, including ones that involve training, can yield new insights. In particular it is useful to see how exposure to words and shapes influences labeling and future word learning.
9. It is reasonable talk about perceptual and cognitive factors as influencing behavior.
10. Experience prior to entering a study can make a difference to patterns of word learning.

#### The curious irrelevance of data

Given such a broad and potentially fertile common ground, where do the parties disagree? They rarely disagree on the data. Indeed, each side doesn’t seem really bothered by the data that the other side presents as seemingly devastating, raising the question as to what data would really be critical for either side and why. They do disagree in terms of general theoretical orientation, sometimes quite profoundly. Again, however, it is less clear how to connect data to these orientations.

Consider first the data. Both Booth and Waxman (2008) and Markson, Diesendruck and Bloom (2008) point to a set of findings that they see as highly problematic for the Attentional Learning Account (ALA): Preverbal infants (as well as very early verbal children) show a shape bias, a shape bias is evident in tasks having nothing

Address for correspondence: Frank C. Keil, Department of Psychology, 2 Hillhouse Avenue, P.O. Box 208 205, Yale University, New Haven, CT 06520-8205, USA; e-mail: frank.keil@yale.edu

to do with naming, the bias can be easily overridden when contexts suggest that it is less relevant or useful, perceived creators' intentions influence the shape bias, causal relations influence the shape bias, and high-level category and functional construals influence the shape bias.

Taken together, these findings are seen as posing a serious challenge to the ALA by undermining its claims that the shape bias is simply a mechanistic association of words and object shapes that can only emerge after considerable experience with word/shape correlations. Yet Samuelson and Horst (2008) and Colunga and Smith (2008) happily accept these findings and don't see them as posing a threat. Indeed, as Colunga and Smith (2008) put it, there is a 'bizarre outcome that proponents of the two sides can conduct nearly identical experiments and each see the same patterns of results as strongly supporting their own position' (p. 203). Somehow, the same studies are being interpreted as having radically different interpretations.

The same phenomenon exists in the other direction. The ALA proponents point to their own empirical findings as critically supportive. For example, they discuss the results of a longitudinal training study and of neural net simulations, results they see as strongly supporting their position that statistical patterns of perceptual inputs drive performance. Yet, Markson *et al.* (2008) and Booth and Waxman (2008) see no problem at all in accepting those patterns of results. They freely acknowledge that prior experiences could influence word learning and the degree to which the shape bias matters, but they do not see such influences as excluding other kinds of influences that they also see as important.

### The heart of the controversy

How could such a large body of empirical evidence not influence one side or the other? All theories tend to treat isolated pieces of discrepant data as minor anomalies, but when there is a systematic and extensive set of studies whose results are largely uncontested and which each side thinks definitely supports their theory, something has gone wrong. The answer seems to revolve around much larger-scale theoretical orientations that act as strikingly different interpretative lenses on these results.

The ALA proponents argue for Dynamic Systems Theory, in which there is a continuity from the simplest actions of perception to conception. This continuity occurs through a common embedding in the processes of perceiving, attending, remembering and acting. Since almost everything we do requires all of these components, the whole dynamic complex must be considered in attempts to explain any patterns of data. Operationally, this ALA approach has focused on correlations between various perceptual attributes as well as correlations among those correlations (which is seen as the great advance over earlier associationist programs). There is also a considerable distrust of 'amodal propositional thought'

in terms of its efficacy in understanding anything about word learning. I think both sides would agree on this characterization of the ALA approach. It is not clear to me, however, why a dynamic systems view, broadly defined, should focus so much on associations. For example, many in the Gibsonian tradition have felt strong links to dynamic systems theory but have not embraced associative approaches. Even if one should see perception and cognition as part of some seamless dynamic system, that system could work in many other ways than as a massive correlation detector.

Colunga and Smith (2008) do offer an olive branch of sorts at the end of their paper where they acknowledge the reality of higher-order cognitive structures, including propositional ones. They see their enterprise, ALA, as looking at the processes that instantiate those structures. More crudely, the message seems to be that there may be a coarse high-level reality to concepts and propositions but that the interesting details must be understood at the level of operations on statistics over perceptual primitives and in terms of properties of dynamic systems. ALA is portrayed as doing all the heavy lifting, and high-level cognition just provides a shallow gloss on what is going on.

The alternative approach, represented by Booth and Waxman (2008) and Markson *et al.* (2008), seems to be more tolerant of a diverse set of mechanisms and of their levels of influence. It freely acknowledges the influence of statistical regularities, but also sees an influence of cognitions about functions, causes, kinds, intentions and goals. It argues that children, and infants, are aware of quite abstract kind categories and have notions about what sorts of properties are important to these categories. Shape clearly plays a role but is not seen as so distinctively special or so closely related to patterns of word usage. This alternative seems to be a more inclusive view that acknowledges a rich and diverse array of processes in both perception and cognition that can have fine-grained influences on how children acquire and use labels. It does not want to see the shape bias as a relatively autonomous process distinct from the ways other properties might influence inferences about kinds. Moreover, it sees conceptual structures and processes as very much involved in explaining the detailed mechanisms of word learning and not simply as high-level glosses on what is going on. It is ironic that an associationist model (implemented in a connectionist network) is used here to argue for a highly local and domain-specific learning system (just words), while the alternative approach is used to argue that word learning is part of a much more general set of learning processes.

How can we choose between these two theoretical approaches? As we have seen, they seem to be largely impervious to data. The ALA approach seems to have changed over the years so that less and less data are capable of refuting it. The alternative approach, by virtue of adopting a more pluralistic perspective, does not seem to have encountered as much of a need to

change. It does seem, however, to be more easily refutable in principle. If, for example, object function, creator's intent, or pre-linguistic behavior had no relation to reliance on shape, that account would be in serious trouble. If shape always trumped other properties and relations, such as functioning as container, that would be a problem as well. In fact, there seem to be lots of uncontested experimental findings supporting such effects and thus not threatening this alternative approach.

In the interests of full disclosure, I will confess to being partial to the pluralistic view (Keil, 1991, 1998; Yamamoto & Keil, 2000). I have been repeatedly impressed by the ways in which abstract conceptual relations seem to influence cognition from infancy onwards (Keil, 2006). I do not, however, see a necessary conflict between such a perspective and all connectionist views. There are, in fact, efforts to explain how parallel distributed processing (PDP) models might yield highly abstract categories and concepts, such as those involving function, that could in turn influence further learning (Rogers & McClelland, 2004). Moreover, those higher-order correlations among correlations can come to have a stronger influence than lower-level ones between perceptual primitives. In essence, PDP systems can be understood as engaging in factor analysis, where the factors or principle components can be quite abstract and come to have a life of their own (Ghahramani & Hinton, 1998).

In the end, these more recent connectionist approaches may not be able to fully account for high-level cognitive processes, but they do seem to at least offer a way of incorporating some of the properties and types often associated with abstract concepts and relations. In principle, why not try to use such approaches to more explicitly model the results of the sort mentioned by Booth and Waxman (2008) and Markson *et al.* (2008)? As best I can tell, there is no reason why PDP approaches need relegate conceptual relations to the realm of only coarse influences and why they could not instead have fine-grained influences very early in development. Similarly, I'm not sure why a dynamic focus on process should necessarily diminish the role of concepts.

More generally, debates over the internal representational formats underlying behaviors have historically proven to be extremely difficult to resolve. Whether a representation is image-like or propositional, exemplar-based or prototype-based, or conceptual vs. perceptual often seems to lead to endless rounds of experimental studies with no resolution. There is another research strategy that may often be a better way to start, one that focuses on the kinds of real-world patterns that seem to drive behavior rather than on the underlying representational format. Here again, the work of the Gibsons (Gibson, 1969; Gibson, 1979) is relevant. For the Gibsons, it was patterns like flow fields and invariants that specified depth. For a child learning word meanings, that information might occur in several forms. It might be causal patterns that link the properties together, such as Ahn's causal status effect (Ahn, Gelman, Amsterlaw, Hohenstein &

Kalish, 2000). It might be logical relations such as transitivity. Or it might be goals and intentions. Most people from all points on the theoretical spectrum think that such patterns are real things that occur in the world. One can then ask which ones influence the behavior in question and in what ways. ALA seems to think that a much smaller range of patterns has most of the influence not only on the shape bias but also more broadly on the acquisition of word meaning. Formulated more precisely, this could be tested against the predictions of the alternative view. It might also lead to decomposition of the meaning of 'shape', that is, whether it really is used in the same way for all artifacts and animals. My guess is that it will not and that aspects of contour, part and pattern shape will have quite different influences as a function of the kind involved (Levin, Takarae, Miner & Keil, 2001). But this question can only be asked by more carefully specifying the kinds of patterns that might be relevant.

We may make real progress on the representational formats that underlie complex behaviors such as word learning, but a necessary first step may be to better understand the kinds of real-world patterns that are tracked by children and how those tracking patterns change over development. To date, these have been studied in an ad hoc manner by adversaries looking for ammunition to marshal against one view or the other. Instead, it would be much more helpful in the future for researchers to systematically lay out the kinds of patterns that might be used by children and then see what they actually do use. Those findings might then be used to build a theory. In doing so, perhaps these different groups would find more harmony among their views.

## Acknowledgement

Preparation of this paper and some of the research described therein was supported by NIH Grant R37HD023922 to F.C. Keil.

## References

- Ahn, W., Gelman, S.A., Amsterlaw, J.A., Hohenstein, J., & Kalish, C.W. (2000). Causal status effect in children's categorization. *Cognition*, **76**, B35–B38.
- Booth, A.E., & Waxman, S.R. (2008). Taking stock as theories of word learning take shape. *Developmental Science*, DOI: 10.1111/j.1467-7687.2007.00664.x
- Colunga, E., & Smith, L.B. (2008). Knowledge embedded in process: the self-organization of skilled noun learning. *Developmental Science*, DOI: 10.1111/j.1467-7687.2007.00665.x
- Ghahramani, Z., & Hinton, G.E. (1998). Hierarchical non-linear factor analysis and topographic maps. In M.I. Jordan, M.J. Kearns, & S.A. Solla (Eds.), *Advances in neural information processing systems 10*, NIPS\*97 (pp. 486–492). Denver, CO, 1–6 December 1997. Cambridge, MA: MIT Press.
- Gibson, E.J. (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.

- Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Keil, F.C. (1991). Theories, concepts, and the acquisition of word meaning. In J.P. Byrnes & S.A. Gelman (Eds.), *Perspectives on language and cognition: Interrelations in development* (pp. 197–224). Cambridge: Cambridge University Press.
- Keil, F.C. (1998). Cognitive science and the origins of thought and knowledge. In R.M. Lerner (Ed.), *Theoretical models of human development*. Volume 1 of W. Damon (Ed.-in-chief) *Handbook of child psychology* (5th edn.). New York: Wiley.
- Keil, F.C. (2006). Cognitive science and cognitive development. In W. Damon & R. Lerner (Series Eds.) & D. Kuhn & R.S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol 2: Cognition, perception, and language* (6th edn., pp. 609–635). New York: Wiley.
- Levin, D.T., Takarae, Y., Miner, A., & Keil, F.C. (2001). Efficient visual search by category: specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception and Psychophysics*, **63**, 676–697.
- Markson, L., Diesendruck, G., & Bloom, P. (2008). The shape of thought. *Developmental Science*, DOI: 10.1111/j.1467-7687.2007.00666.x
- Rogers, T.T., & McClelland, J.L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.
- Samuelson, L.K., & Bloom, P. (2008). The shape of controversy: what counts as an explanation of development? Introduction to the special section. *Developmental Science*, DOI: 10.1111/j.1467-7687.2007.00663.x
- Yamamoto, K., & Keil, F.C. (2000). The acquisition of Japanese numerical classifiers – linkages between grammatical forms and conceptual categories. *Journal of East Asian Linguistics*, **9**, 379–409.

DOI: 10.1111/j.1467-7687.2007.00669.x

## The shape bias: an important piece in a bigger puzzle

Jeffrey L. Elman

Department of Cognitive Science, University of California, San Diego, USA

### Introduction

No theory remains fixed for ever, at least not until it is abandoned. Theories develop over time. Some changes reflect the discovery of new facts that prompt modification of the earlier version. And often, especially in the case of theories that are tied to mechanisms, changes arise because with time comes a deeper understanding of the mechanism's characteristics. This understanding in turn generates greater precision in the predictions that can be made by the theory. Obviously, a critical issue is whether the changes are post hoc and unprincipled – in which case the theory is at risk of losing credibility – or whether the developments reflect true maturation and greater insight.

I believe that changes in the evolution of the Attentional Learning Account (ALA) play a large role in the apparent disagreements that motivate the current set of papers. In fact, there is considerably more agreement than appears on a casual reading of these papers. But I also believe there remain significant points of disagreement. In this commentary, I will try to address what I see as the issues on which there is convergence, and those for which there remain important differences.

### Apparent disagreements

Let me address what seem to be the major disagreements. These revolve mainly around the following claims:

1. The shape bias is language/naming specific. It does not appear in non-linguistic domains.
2. The shape bias only emerges when language learning is well on its way, typically at a stage where a child has a productive vocabulary of 50–150 nouns.
3. The shape bias is sensitive only to perceptual (not conceptual) factors.

For each of these, there appear to be two camps. One camp says Yes to each of these propositions, and is represented in the papers by Colunga and Smith (2008; CS), and by Samuelson and Horst (2008; SH). This is the ALA group. The other camp says No, and is represented by Booth and Waxman (2008; BW), and by Markson, Diesendruck and Bloom (2008; MDB). This is the anti-ALA group. To be clear: I do not claim that any of these authors would be happy with this heavy-handed parsing of positions. Indeed, I am quite sure that several of them would object, particularly given the simplistic way I've phrased things. But it provides a start.

Address for correspondence: Jeffrey L. Elman, Department of Cognitive Science, University of California, San Diego, La Jolla, CA 92093-0515, USA; e-mail: jelman@ucsd.edu