SHAPE BIAS SPECIAL SECTION

Confronting complexity: insights from the details of behavior over multiple timescales

Larissa K. Samuelson¹ and Jessica S. Horst²

1. Department of Psychology and Iowa Center for Developmental and Learning Sciences, University of Iowa, USA
2. Department of Psychology, University of Iowa, USA

Abstract

Young children tend to generalize novel names for novel solid objects by similarity in shape, a phenomenon dubbed ‘the shape bias’. We believe that the critical insights needed to explain the shape bias in particular, and cognitive development more generally, come from Dynamic Systems Theory. We present two examples of recent work focusing on the real-time decision processes that underlie performance in the tasks used to measure the shape bias. We show how this work, and the dynamic systems perspective, sheds light on the controversy over the origins and development of the shape bias. In addition, we suggest that this dynamic systems perspective provides the right level for explanations of development because it requires a focus on the details of behavior over multiple timescales.

Introduction

Young children learn new nouns at astonishing rates. In the laboratory, they are able to generalize novel nouns to new instances after seeing a single instance named once. For example, when shown a novel solid object named with a count noun, young children tend to generalize the novel name to new instances that match in shape, a phenomenon dubbed the ‘shape bias’. Since Landau, Smith and Jones’ (1988) seminal study on this bias, it has been replicated in numerous laboratories as is reflected in the four papers that serve as the focus of this special section. Two- and 3-year-old children generalized novel names for novel solid objects by shape in Diesendruck and Bloom’s (2003) studies, as did 2-year-old children in Booth, Waxman and Huang’s (2005) study. Similarly, following 9 weeks of special category training, the 15–21-month-old children in Smith, Jones, Landau, Gershkoff-Stowe and Samuelson’s (2002) and Samuelson’s (2002) studies demonstrated a shape bias. Thus, there is no dispute among these papers regarding the behavior – young children attend to shape when naming novel solid objects. The question is what does this behavior mean, and where does it come from? At this level, the devil is in the details.

In their longitudinal training studies, Samuelson (2002) and Smith et al. (2002) found that whether children developed a shape bias depended on exposure to a noun vocabulary dominated by names for solid things in categories organized by similarity in shape. Thus, these researchers argue that the noun vocabulary children learn early teaches them what features of objects to attend to when learning novel names. In contrast, Diesendruck and Bloom (2003) found no differences in children’s generalizations when asked to pick objects that were ‘of the same kind’ or to generalize novel names. They also failed to find differences in the level of 2- and 3-year-old children’s attention to shape. Thus, Diesendruck and Bloom (2003) argue that the shape bias reflects children’s knowledge that shape is a good indicator of object kind. Similarly, Booth, Waxman and Huang (2005) found that children’s tendency to generalize a novel name based on shape similarity depended on what they were told about the exemplar – even for very young language learners. They thus argue that the shape bias is based on conceptual knowledge that does not develop as a result of early vocabulary learning.

We contend that the critical insight needed to understand the details of when children attend to shape and what it means comes from Dynamic Systems Theory. In particular, making progress in understanding the development of word learning biases and cognitive development more generally requires understanding how children’s knowledge about names and categories is brought to bear in a task in a moment in time, and how individual behaviors at that timescale accumulate to create later behaviors. In what follows, we present examples of two recent lines of work exploring the time-dependent processes that create novel noun generalization behaviors. These examples demonstrate the importance of understanding behavioral processes at these timescales for explanations of the origins of the shape bias. We illustrate...
this point by showing how a dynamic systems perspective leads to a reinterpretation of key results that form the basis of the current debate on the shape bias. We conclude by suggesting that explanations of the shape bias, and cognitive development more generally, need to take these time-dependent influences on behavior into account.

A dynamic systems account of the shape bias

Our dynamic systems approach to the origin of the shape bias extends the Attentional Learning Account (ALA) proposed by Smith and colleagues (see Smith & Samuelson, 2006). Briefly, the ALA proposes that as children learn their early vocabulary, they learn a system of statistical regularities among linguistic devices, the properties of objects, and perceptual category organization. These learned associations then mechanistically shift attention to the relevant properties of objects in future word learning situations, enabling children to make generalizations from the categories they know to novel categories. As a consequence, children learn new words rapidly. By this account, children's attention to shape results from the particular statistical structure of the linguistic context they have previously been exposed to. Thus, the ALA predicts individual, cross-cultural, and task-based differences in performance that are closely tied to the specifics of an individual's history, the particular language being learned, and the specifics of the task – predictions supported in a growing body of work (see Smith & Samuelson, 2006, for a review). Simulation studies with connectionist networks have been used to formally instantiate these ideas (Colunga & Smith, 2005; Samuelson, 2002). Further, the links between the vocabulary, the development of biases, and subsequent word learning have been confirmed in longitudinal training studies in which young children taught a noun vocabulary dominated by names for categories organized by shape, but not children taught vocabularies organized in other ways, subsequently demonstrate precocious biases and an acceleration in vocabulary development (Samuelson, 2002; Smith et al., 2002).

The ALA itself stems from a more general dynamic systems perspective of development that, rather than seeing knowledge as conceptual information tapped into via a perceptual system, views it as the emergent product of the real-time interaction of many components – the child's physical and cognitive abilities, the specifics of the task, the state of the system in the just-previous past, and the child's individual developmental history of perceiving and acting (Smith & Thelen, 2003; Thelen & Smith, 1994). By the dynamic systems perspective, each individual moment of knowing reflects the context-specific integration of these components and sets the stage for the next interaction between the child and the context. Over time, stable patterns of behavior emerge as many individual interactions are laid down over the course of a child's developmental history. Likewise, variability and change emerge as the system is pushed into new forms of organization by the specifics of individual contexts, abilities, and experiences. Thus, the dynamic systems view accounts for both stability and variability in performance, grounds current behavior in the longer-term activity of the system, and is fundamentally developmental (see also Samuelson & Smith, 2000b).

Many connectionist approaches share a similar emphasis on development (see Spencer, Thomas & McClelland, forthcoming; Smith & Samuelson, 2003; Thelen & Bates, 2003); thus, it should not be surprising that the ALA has previously been implemented in connectionist models (Colunga & Smith, 2005; Samuelson, 2002). Importantly, however, these formal instantiations of the ALA do not specify how decision processes unfold in real-time. Our recent work has extended the ALA in this new direction using a new class of dynamic systems models – dynamic fields (Spencer & Schöner, 2003). In addition, we have begun to investigate the role of children's just-previous history in shaping behavior, inspired by previous work using dynamic fields to capture the A-not-B error. We present each of these examples in turn.

Decisions in real-time

As a first step towards a more complete understanding of the processes that create word learning biases, we have been investigating how children make decisions in noun generalization tasks in real-time. This critical determinant of how children behave in novel noun generalization tasks has received relatively little attention in the field (but see Deák & Bauer, 1995, for relevant work). The two most common tasks used to elicit novel noun generalizations from young children are yes/no procedures and forced-choice procedures. In both tasks, children are shown a novel exemplar object which is named. In yes/no tasks, children are then shown each test object individually and asked whether each can be called by the same name as the exemplar (e.g. Samuelson & Smith, 2000a). In forced-choice tasks, however, children are shown the test objects simultaneously and asked to pick which one can be called by the same name as the exemplar (e.g. Diesendruck & Bloom, 2003). We have used an extension of the Dynamic Field Theory (DFT) to simulate the decision processes that underlie performance in these tasks (Samuelson, Horst, Dobbertin & Schutte, 2006) because the DFT has been used to capture the real-time processes that underlie how children make and maintain stable decisions in working memory (Simmering, Spencer & Schöner, 2006; Spencer, Simmering & Schutte, 2006). This is a novel extension of the DFT (but see Lipinski, Spencer, Samuelson & Schöner, 2006; Smith & Samuelson, 2005, for related ideas), that explicitly draws on what is known about basic perceptual and working memory processes to explore the generation of a real-time decision in novel noun generalization tasks. This work demonstrates the importance of this timescale for our understanding.
of children's novel noun generalization behaviors and lays the groundwork for future explorations of the relation between real-time decisions and developments at the next timescale—the accumulation of statistical regularities that are the heart of the ALA.

Figure 1 shows example model simulations that explore the nature of children's decisions as dictated by the developmental state of the child and as constrained by the details of the task and input (see Samuelson, Schutte & Horst, 2007, for details of the model). The top panels in each box depict input to the model in either a yes/no (top and middle boxes) or forced-choice (bottom box) task. The bottom two panels in each box depict younger and older decision fields, that is, fields that capture the decisions children make in these different tasks. The x-axis in each panel shows a set of neurons arranged by ‘similarity’. Neurons that are tuned to respond, for instance, to similar perceptual features would be close neighbors along this dimension, while neurons that are tuned to respond to dissimilar features would be far from one another. The activation of each neuron is plotted along the y-axis. Time is shown along the z-axis as the sequence of events in a single trial unfold (starting at the back). In these simulations, developmental differences between younger and older language learners have been implemented as a change in the precision of neural interactions in the decision field, such that younger language learners have broader, weaker interactions (e.g. panel B) while those of older language learners are narrower and stronger (e.g. panel C; for related ideas, see Schutte, Spencer & Schöner, 2003).

The stimuli are presented to the model via the input field. Inputs take the form of Gaussians representing the exemplar (leftmost peak in panels A, D and G) and either one or two test objects depending on the task. Inputs are positioned along the similarity dimension according to their relative similarities. In the depicted simulations, the shape-matching test object is closer to the exemplar than the material-matching test object. These positions are based on our recent work quantitatively modeling noun generalization data from 30-month-old children in these tasks (Samuelson et al., 2007). Note that the activation of the exemplar in the forced-choice task is lower because it is typically farther from the child, as are all the stimuli in the yes/no task.

The time-dependent processes (i.e. the dynamics) that determine the model's response occur in the decision field. Neurons in this model interact according to a local excitation/lateral inhibition function. This means that neurons with similar 'preferred' inputs excite one another while neurons tuned to very different inputs inhibit each other. This allows the network to form stable peaks of activation that represent behavioral decisions to, for instance, select a particular input in a forced-choice task. Task differences emerge from differences in the strength and time structure of the inputs and the dynamics of the decision field. The yes/no task presents a less constrained type of decision because a 'yes' response can be generated.

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd.
to multiple test objects as long as each is relatively similar to the exemplar. This is captured in the model by making the neurons in the decision field more excitable and less competitive, thereby allowing more blending across stimuli and good sensitivity to overlap in the input. The formation of a stable peak in the decision field represents a ‘yes’ response, and the failure to form a stable peak, a ‘no’ response (Simmering et al., 2006).

In contrast, the forced-choice task presents a more constrained type of decision because the child must pick one test item. Such selectivity is achieved by making the decision field highly competitive (i.e. the neurons interact strongly) and making it reflect the inputs strongly (Willmzig & Schöner, 2005). At the point in the trial when the test objects are moved within reach of the child and a response is requested, the resting level of neurons is raised. In the model, this change further increases competition among the neurons and the overall activation in the field, thereby forcing the creation of a single, stable peak centered over one of the test objects. The model’s choice corresponds to the location of this peak along the x-axis.

As can be seen in Figure 1, the combination of the developmental differences and task dynamics leads to very different patterns of performance – even given the same inputs. In the yes/no task (top and middle boxes), the model, like children, is presented with each input separately. Thus, the two peaks in panel A correspond to the exemplar and the shape-matching test object, while the peaks in panel D correspond to the exemplar and the material-matching test object. Panel B shows the result of presenting the shape-matching test object to the younger model – a peak forms indicating a ‘yes’ response. Likewise, the older model generates a ‘yes’ response to the same test object (panel C). The younger model also generates a ‘yes’ when the material-matching test object is presented (panel E), whereas the older model fails to form a peak and generates a ‘no’ response (panel F). Thus, the narrower interactions in the older model lead to greater selectivity – a bias to say ‘yes’ only to shape-matching stimuli.

In the forced-choice task (bottom box), the model, like children, is presented with an exemplar plus two test objects at the same time (panel G). Note that the positions of the stimuli along the similarity dimension and their precision (i.e. the widths of the inputs) are the same as those presented in the yes/no task (panels A and D). As can be seen in the figure, both the younger (panel H) and the older (panel I) versions of the model formed a peak centered at the input location of the shape-matching test object.

Importantly, developmental changes were instantiated in the same way for simulations of both the forced-choice and yes/no tasks. Thus, these example simulations demonstrate that the specific question posed by the task matters. The yes/no task is less constrained; thus, to get generalization that is specific to one object but not the other requires changes in the decision-making processes themselves. Conversely, the performance of the forced-choice models suggests that a lack of differences across age in this task may mask changes in the underlying processes. We reiterate that we have used this model to quantitatively capture differences in children’s performance in these two tasks. This work creates an important bridge between word learning and other areas of cognitive development in that it ties the decision processes underlying performance in the novel noun generalization tasks to decision-making in other cognitive tasks.

Decisions make history

Our next step towards a more complete understanding of the processes that create word learning biases was to investigate how individual decisions made in real-time accumulate to influence behavior at the next timescale. Research with infants on Piaget’s classic A-not-B task demonstrates links between these timescales. In this task, a toy is hidden at an A location and the infant is allowed to search and find the toy. This is repeated several times before the experimenter switches and hides the toy at a B location. Even after watching the toy being hidden at the new location, 8- to 10-month-old infants err on this B trial and search for the toy at A. Thelen, Smith and their colleagues have proposed that the source of infants’ error in this task is the history of reaching to A repeatedly in the context of a relatively weak ability to maintain a decision to reach to B (Smith, Thelen, Titzer & McLin, 1999). A dynamic field model provides the mechanism: A trace of activity is laid down by each behavioral decision to reach a particular location. These traces accumulate across the A trials to influence later behavior. Recent research in our laboratory suggests that a similar process may influence children’s novel noun generalization behaviors over the course of an experimental session (Samuelson & Horst, 2007).

We presented 24-month-old children with either a novel solid or nonsolid exemplar that was named. In a forced-choice task, children generalized the name to either a shape-matching or a material-matching test object. Our interest was in the aspects of the task and stimulus context that push children to attend to shape when generalizing names for solid objects but material when generalizing names for nonsolid substances. Across four conditions, we systematically manipulated three aspects of the task that appeared to be linked to differences in attention to shape or material in prior studies (see Samuelson & Smith, 1999; Soja, Carey & Spelke, 1991) – the training children received, whether solid exemplars were always presented as whole objects or in pieces on some trials, and whether material-matching test objects also matched exemplars in color. All the children had at least 150 nouns in their productive vocabulary – a level previously shown to be associated with significant attention to shape in tasks that include both solid objects and nonsolid substances (Samuelson & Smith, 1999).

As can be seen in Figure 2, the biases children demonstrated were influenced by the training and the specifics
of the stimuli. Children in the Solid-Only+Wholes+Material-Only condition (S+W+M) received two training trials with solid objects, saw only whole solid objects on the test trials, and saw material-matching test objects that did not match the exemplar in color. As in previous studies with similar conditions (i.e. Samuelson & Smith, 1999), the proportion of shape choices in this condition was significantly greater than levels expected by chance on trials with solid exemplars, but not different from chance on trials with nonsolid substances. These children demonstrated a significant shape bias with solid exemplars but no significant ‘material bias’ with nonsolid exemplars.

When some of the solid exemplars were presented in pieces but the same training regime and material-only-matching test objects were used (S+P+M), children still demonstrated a significant shape bias and no significant material bias. However, when the training regime was changed to include both solid objects and nonsolid substances (SN+P+M), children no longer demonstrated significant biases. Finally, when material-matching test objects also matched the exemplars in color (SN+P+MC), children demonstrated a significant material bias, as in prior studies that used material+color-matching test objects (e.g. Soja et al., 1991).

Thus, like studies of the A-not-B error, these results suggest that the training trials critically influence what children do later in the experiment. A further analysis of the data suggests other effects at the timescale of trial-by-trial performance. The left panel of Figure 3 presents the percentage of children from the three conditions that included exemplars in pieces (all but S+W+M) who chose shape-matching test objects on their first trial with a solid-whole-exemplar (SW1) as a function of how many trials with exemplars (solid or nonsolid) in pieces they had previously seen. As can be seen in the figure, children generalized the novel name for solid whole exemplars by shape similarity if they had not previously seen any exemplars in pieces. However, as children saw more exemplars in pieces over the course of the experiment, they were less likely to generalize novel names for solid-whole exemplars by shape similarity. Further, children’s first generalizations of a name for a solid-whole exemplar influenced their subsequent generalizations. The right panel of Figure 3 presents performance on the second trial with a solid-whole exemplar (SW2) as a function of whether individual children had generalized the novel name on SW1 by shape or material similarity. As can be seen in the figure, children who generalized the novel name by shape on SW1 were more likely to also generalize the novel name by shape on SW2.

These data fit with a dynamic systems perspective by showing how the trial-to-trial sequence of events (how many times a child sees exemplars in pieces) and the child’s active responses (having generalized a name for a solid object by shape recently) build on themselves to create short-term statistical biases to generalize novel names by shape or material in an experiment. These effects will require further investigation via targeted manipulations of trial orders and training sequences. Nevertheless, they are some of the first to make the link between the timescales of real-time decisions and later behaviors (see also Landau & Shipley, 2001).

The recent work from our lab combined with previous data showing that children given experience with biased...
short-term statistics develop a precocious shape bias and accelerated vocabulary development (Samuelson, 2002; Smith et al., 2002) begin to provide a bridge between the timescale of moment-to-moment behavior and the longer timescale at which word learning and development emerge. This work also sheds light on the debate over what children’s biased noun generalization behavior means and where it comes from because it provides insight into some of the differences in the details of behavior seen in the four papers that are the focus of this special section. For example, Diesendruck and Bloom (2003) argued against the idea that the shape bias develops based on the fact that there was no difference in the level of shape bias demonstrated by 2- and 3-year-old children. However, as the task analysis and model presented above suggest, it is possible that the use of a forced-choice task in Diesendruck and Bloom’s studies may have masked underlying differences in children’s understanding of the importance of shape for nominal categories. That is, it is possible that the 3-year-old children in their study did have a greater appreciation than the 2-year-old children regarding the importance of shape for categories of solid things. However, because the statistics of both groups’ vocabularies directed attention to shape with solid objects, both groups were more likely to pick the shape-matching test object over the material-matching one in a forced-choice task that maximizes even subtle differences between stimuli.

Similarly, Booth, Waxman and Huang (2005) found that 18-month-old children demonstrated a shape bias, and found no difference in their responses when compared to a group of 24-month-old children tested with the same procedure. Thus, they argued that the shape bias is based on conceptual knowledge that does not develop as a result of early vocabulary learning. However, as they note, getting very young children to respond in noun generalization tasks requires extra training. Thus, prior to the experimental trials, children in their task were introduced to two distracters and explicitly told that they were not called by the same name as the previously introduced exemplar. Then, an object that was identical to the exemplar was introduced and children were told that it was called by the same name as the exemplar. In contrast, the studies that have suggested that the shape bias develops later (Gershkoff-Stowe & Smith, 2004; Samuelson & Smith, 1999) have used a task that required children to generalize a novel name after a single presentation of a single exemplar (and no contrast items). In light of our data suggesting that the specifics of the task and the training children receive can have a profound influence on the behavior they demonstrate, it is not clear that Booth et al.’s data create a conflict. Rather, they reinforce the idea that task-specific details matter, showing that young children can do something in a supportive task that they cannot do in a less supportive task. Indeed, our dynamic systems perspective suggests that understanding the link between the details of the task and the specific behaviors produced will provide important insight into the processes underlying novel noun generalization.

It is important to note that we are not trying to dismiss or trivialize the data from these studies or differences between studies. Rather, we argue that it is exactly such details of behavior over time that matter to all explanations of development. More specifically, any explanation of development should include an appreciation of three critical issues. First, empirical details matter because it is the details of the interaction between the child and the task that make behavior. From this perspective, formal models are useful because they necessitate specificity about the nature of the task, the stimuli, the knowledge brought by the child, the underlying assumptions, and how all these components interact. Second, our means of assessing mental states is always via behavior (see also Bremer & Mareschal, 2004). Thus, a rich understanding of behavior is essential if we are to understand underlying mental states. Third, development is created over multiple timescales. A full explanation of the origins of any behavior, therefore, will require understanding how individual behaviors in a moment accumulate to create longer-term patterns in development.

This is why, by our view, explanations at the level of ‘kinds’ or ‘conceptual understanding’ are not satisfying – they lack clarity with regard to these critical elements. Our discussion above illustrates how these explanations can be augmented by considering the details of how the child and task interact to create the mental states that underlie behavior. But clarity regarding the third element is perhaps most critical – the creation of development over time. Explaining the shape bias as based on an understanding of ‘kinds’ or ‘conceptual understanding’ begs the question of where an understanding of kinds or concepts comes from. Booth, Waxman and Huang (2005) point to this issue when they state that ‘conceptual understanding’ may be learned via associative mechanisms similar to the ALA (p. 502). Hence, even explanations based on higher level concepts must specify the nature of learning, how what is learned is brought to bear in specific task contexts, and how learning accumulates over time to create stable patterns of behavior. In short, a full explanation of the shape bias, or any other cognitive phenomenon, will require attention to precisely those details of behavior in time that we have highlighted here. Clearly, we have taken only the first steps in the direction of this more complete account. Nevertheless, as our examples illustrate, these are critical steps to take in order to resolve the debate on the origins of the shape bias, and for explanations of development more generally.

Acknowledgements

Preparation of this manuscript was supported by NICHD grant number 5 R01 HD045713 to the first author. The authors wish to thank John Spencer, Jodie Plumert and Linda Smith for thoughtful discussions of this work and John Spencer for helpful comments on earlier versions of the manuscript. We also thank Ryan Brink, Brandi...
Dobbertin, Anne Schutte and the members of the Language and Category Development Lab for their assistance. Finally, we thank the parents and children who participated in our the studies, without whom this work would not be possible.

References


