An Attentional Learning Account of the Shape Bias: Reply to Cimpian and Markman (2005) and Booth, Waxman, and Huang (2005)

Linda B. Smith  
Indiana University

Larissa Samuelson  
University of Iowa

Recently, Developmental Psychology published two articles on the shape bias; both rejected the authors’ previous proposals about the role of attentional learning in the development of a shape bias in object name learning. A. Cimpian and E. Markman (2005) did so by arguing that the shape bias does not exist but is an experimental artifact. A. E. Booth, S. R. Waxman, and Y. T. Huang (2005), in contrast, concluded that the shape bias (and its contextual link to artifact categories) does exist but that the mechanisms that underlie it are conceptual knowledge and not attentional learning. In response, in this article the authors clarify the claims of the Attentional Learning Account (ALA) and interpretations of the data under question. The authors also seek to make explicit the deeper theoretical divide: cognition as sequestered from processes of perceiving and acting versus as embedded in, and inseparable from, those very processes.

Keywords: word learning, vocabulary development, shape bias

Developmental Psychology recently published two articles on the shape bias in early noun learning that conclude, on very different grounds, that our account of the phenomena is wrong. We were not called upon to comment on these articles in the review process, and thus we do so now. Our goals are to clarify our account and open discussion of several issues in this contentious literature.

The Attentional Learning Account (ALA) of Early Noun Learning

The past 20 years of research on early noun learning provide a well-documented set of to-be-explained phenomena. In brief, the data are these: One-year-olds typically comprehend and produce a few object names. Two-year-olds, in contrast, comprehend and produce several hundred object names (e.g., Fenson, Dale, Reznick, & Bates, 1994). Moreover, in experimental tasks, children as young as 2–2 1/2 years of age show systematic attentional biases, exquisitely attending to different properties for different kinds—to the multiple similarities of animals, the shapes of artifacts, the material of substances, or the colors of foods (e.g., Booth & Waxman, 2002; Imai & Gentner, 1997; Jones & Smith, 2002; Macario, 1991; Soja, Carey, & Spelke, 1991).

Attention to the shape of artifacts is the most well-studied of these attentional biases. This “shape bias” is evident in quite young word learners but with age becomes more robust and more specific to things with the perceptual properties of artifacts (e.g., Colunga & Smith, 2005; Gershkoff-Stowe & Smith, 2004; Graham & Poulin-Dubois, 1999; Hupp, 2004; Jones, Smith, & Landau, 1991; Landau, Smith, & Jones, 1988) and to linguistic contexts indicative of artifacts rather than of substances or animals (Booth & Waxman, 2002; Gathercole, 1997; Soja, 1992; Yoshida & Smith, 2001, 2003a). The shape bias is evident in children learning different languages, but the developmental trajectory and relevant contextual cues (both perceptual and linguistic) vary across languages (Colunga & Smith, 2005; Gathercole, Thomas, & Evans, 2000; Imai & Gentner, 1997; Yoshida & Smith, 2003a). Finally, biased attention to shape is weaker in children with language delay (Jones, 2003; Jones & Smith, 2005) and may be an early warning of a potentially significant developmental disorder (Rescorla, Roberts & Dahlsgaard, 1997; Thal & Katzic, 1996).

The ALA explains this full array of results through three major claims. The first claim is that the learning environment presents correlations among linguistic devices, object properties, and perceptual category organization. We have provided support for this idea by studying the statistical structure of the first 300 nouns in English and in Japanese (Colunga & Smith, 2005; Jones & Smith, 2002; Samuelson & Smith, 1999; Smith, Colunga, & Yoshida, 2003; see, also Sandhofer, Smith, & Luo, 2000). Among early learned noun categories, artifacts tend to be rigid, angular, solid things in categories organized by shape. Animals have eyes, legs, and heads and are in categories organized by multiple similarities. Substances tend to be nonsolid and in categories organized by material. Thus, as illustrated in Table 1, correlations can be seen between the perceptual properties of individual things and the kinds of similarities characteristic of instances in a common noun category. Critically, languages also present other words (determiners, classifiers, verbs) that are, because of their use with nouns of different kinds, also correlated with perceptual properties and with the similarities characteristic of instances of a noun category.

The second claim of the ALA is that as children learn lexical categories, they learn these statistical regularities and the higher order generalizations that emerge from them, enabling the even more rapid learning of lexical categories. Support for this developmental claim lies in a wide range of studies (for reviews, see Smith, 1999, and Smith, Colunga, & Yoshida, 2003). The stron-
gest evidence derives from (a) longitudinal studies that showed a
tight temporal link between the particular nouns that an individual
child knows and that child’s attentional biases in artificial noun-
learning tasks and (b) microgenetic studies in which vocabulary
growth was experimentally accelerated. In one 6-month longitu-
dinal study of children’s acquisition of their first 100 nouns,
Gershkoff-Stowe and Smith (2004) showed that the emergence of
a robust shape bias in individual children coincided with an ac-
celeration in that child’s nominal vocabulary growth. In training
experiments (Samuelson, 2002; Smith, Jones, Landau, Gershkoff-
Stowe, & Samuelson, 2002), 15- to 19-month-olds were taught
lexical categories well organized by shape that they generalized to
novel categories. This training also caused an acceleration (relative
to children in control conditions) in the rate of real-word vocab-
ulary growth outside the laboratory. These longitudinal and mi-
rogenetic studies demonstrated that attention to object shape in
learning names for artifacts is relevant to the real-life vocabulary
development of young children.

The third claim is that children’s learning of the statistical
regularities and their application of that learning in the real-time
task of generalizing a name to a new instance is mechanistically
realized through learned associations that yield contextually cued
shifts in attention. According to ALA, the system of learned
associations illustrated in Table 1 creates contextually cued atten-
tional biases. Children’s attention is automatically directed, in the
moment of learning, to similarities that have been systematically
relevant in those linguistic and perceptual contexts in the child’s
past. The core mechanism, then, is the top-down control of atten-
tion by past knowledge (see especially Smith, 2000, 2003, 2005;
Yoshida & Smith, 2003a, 2003b). The claim is that past learning,
activated by cues in the context, modulates attention and, because
these associations are overlearned, does so rapidly through auto-
matic (rather than deliberative or executive) processes (see Smith,
Jones, & Landau, 1996). In this way, attention is dynamically
bound to context, enabling the child to attend to the correct
similarities for the linguistic and perceptual task at hand. This is a
potentially powerful learning mechanism because (a) it is exquis-
itely tied to and integrates multiple (perceptual and linguistic)
contextual cues in the moment; (b) it enables the learner to attend
to the same perceptual object in different ways depending on
context (e.g., “a muffin” cues attention to shape but “some muffin”
cues attention to material); and (c) attention and learning in the
moment are strongly guided by the history of regularities in the
learner’s past. We have provided support for this mechanistic
account in formal models of the attentional learning process and
experimental tests of novel predictions from those simulation
studies (Colunga & Smith, 2003, 2005; Samuelson, 2002; Smith,
1995) through demonstrations of word, property, and task context
effects (Yoshida & Smith, 2003a, 2003b, 2005), through cross-
linguistic studies of children with different developmental histories
(Yoshida & Smith, 2001, 2003a), and through experimental train-
ing studies in which we teach new context cues and show that they
modulate attention in noun learning tasks (Samuelson, 2002; Smith
et al., 2002; Yoshida & Smith, 2005).

Cimpian and Markman (2005): Is There Really
a Shape Bias?

Cimpian and Markman (2005) acknowledge the evidence showing
a shape bias, but they argue against the conclusion that children
“understand” names as applying to “classes of same-shaped things” (p. 1004). Instead, the authors propose that children under-
stand names as being about “kinds.” Cimpian and Markman
(2005) provide evidence in experiments in which 3- to 5-year-old
children were asked to learn new names for instances of well-
known categories. For example, the child might be shown a
detailed picture of a square-shaped purse (with handle, clasp, etc.,
clearly depicted) and told that in “frog talk” it is called a “wug.”
The child was then asked which of the following was also a “wug”: a
bedroom dresser (same square shape), a saddlebag purse (same
kind), or a paper bag (same superordinate category). Cimpian and
Markman (2005) found that in some conditions, children chose
the shape match, but when contextual information was supplied or
when children were provided the choice of saying that the name
did not extend to any of the alternatives, they did not systemati-
cally choose the shape alternative. Cimpian and Markman (2005)
take these results to indicate that the shape bias “is not a genuine
word learning strategy” (p. 1004). We strongly disagree. Three-to
5-year-olds learning “frog talk” for already known lexical catego-
ries is not “genuine” word learning in any sense. In contrast, the
acquisition of the first 100 nouns by 15- to 24-month-olds
(Gershkoff-Stowe & Smith, 2004), the acceleration in the vocab-
ulary growth of 17-month-olds (Samuelson, 2002; Smith et al,
2002), and the struggle of late talkers to add new nouns to their
productive vocabulary (Jones, 2003) are instances of genuine word
learning and, in these cases, attention to shape matters.

Further, the ALA does not claim that children always or only
attend to shape, even for artifacts, and especially not for well-
known artifact categories. In the General Discussion of a prior
article on the ALA (Yoshida & Smith, 2003c), we predicted
Cimpian and Markman’s (2005) results. As young children learn
names for specific categories, they learn the many specific prop-
erties relevant to those categories; for purses, those characteristic
properties may well include aspects of shape, material, latches,
handles, having money inside, and so forth. Instances of basic-level categories typically have much more in common than merely shape (Cree & McRae, 2003; Rosch & Mervis, 1975; Samuelson & Smith, 1999). Moreover, the similarities that characterize any category are statistical, not necessary and sufficient. Thus, to use an example from one of our previous articles, most pickles may be green, bumpy, and pickle shaped, but not all pickles have all these properties. If children are associative learners, they should learn all the statistical regularities that characterize individual categories, and these learned associations should modulate attention. But—and this is the key claim—children also should learn higher (second, third) order correlations that arise over the learned correlational patterns of many different categories. The shape bias is a higher order correlation. It is not a truth about a specific category but a regularity that emerges in a system of many categories.

In brief, children have category knowledge at multiple levels of abstraction. They should know the specific similarities that matter for specific categories (e.g., that pickles are pickle shaped, green, and bumpy) and the higher order regularities that should guide learning, even in cases of novel things about which children have little knowledge. To demonstrate this category knowledge, Yoshida and Smith (2003c) presented children with novel three-dimensional things that were not recognizable instances of any category known by the child. When the item was labeled with a well-known name (e.g., pickle), children attended to the properties relevant to that particular category (e.g., pickle directed attention to color as well as shape). When the item was labeled with a novel name, children attended exclusively to shape. According to the ALA, this is because the word “pickle” is a context cue that activates a web of specific associations about pickles and, in so doing, shifts attention weights to the properties that have been relevant to pickles in the past. In contrast, the novel noun “wug,” and the syntactic and social context in which it is embedded, is associated with (and, thus, activates, to some degree) all known nominal categories and thus directs attention to the similarity most pervasively and systematically relevant to all known noun categories—shape. Likewise, the ALA predicts that the highly recognizable pictures used by Cimpian and Markman (2005) should activate children’s specific knowledge about specific categories and should direct attention to the cluster of properties relevant to that particular category, not just shape.

In sum, the ALA does not claim that children only and always attend to shape, nor does it claim that they know nothing about other properties that are correlated with specific lexical categories. Rather, the ALA claims that children learn predictive relations between perceptible cues (words, properties, actions) and category organization and that they learn those associations at lower and higher orders of generalization.

Booth, Waxman, and Huang (2005): Conception, Not Perception

Booth et al. (2005) characterize the current state of the field as follows:

Most contemporary theories of word learning reflect this state of the evidence, claiming that in the process of word learning, infants weave together information from various sources, including perceptual, social, pragmatic, and conceptual resources . . . However, this view has not been unilaterally [sic] endorsed. On the contrary, in an influential ongoing program of research, Linda Smith, Susan Jones, and their colleagues have argued that word learning “primarily engages the perceptual systems and is immune to influences from general world-knowledge.” (p. 492; Landau, Smith, & Jones, 1998, p. 20)

This nearly decade-old quote was snipped from context, distorting its meaning and misrepresenting our position. The full quote in the first sentence of the General Discussion was not a summary of our position or data but, rather, was offered as a strong benchmark against which to discuss experimental results:

The present studies sought to determine whether early object naming primarily engages the perceptual systems and is immune to influences from general world-knowledge, and if so, how this might change developmentally. (Landau et al., 1998, p. 20)

Booth et al. (2005) sought to show that more than perceptual information is relevant to children’s noun extensions. In so doing, they concentrated on our previous findings, which showed that (a) the addition of eyes (or legs or shoes) to the named object shifts children’s attention to multiple similarities (e.g., to texture jointly with shape) and (b) that this effect of eyes (or legs or shoes) on children’s formation of lexical categories increases with age and, specifically, with the number of animal names in children’s vocabularies (Jones & Smith, 2002; Yoshida & Smith, 2001). The ALA explains these results in terms of learned associations between object properties and the perceptual similarities that characterize members of the same lexical category and the higher order generalizations that such a system of associations yields.

Booth et al. (2005) reported a set of experiments showing that the words experimenters used to describe objects (e.g., saying, that the object is “happy”) change children’s name extensions so that the presence of the words alone (no eyes, no legs) can bring about patterns of name extensions typical of animal categories, and the authors demonstrated this effect in children as young as 18 months. We like the task and the results and have conducted nearly identical experiments showing that we can alter children’s name extensions by merely changing the words that we say (Yoshida & Smith, 2003a, 2003b). Words (linguistic cues) have played an important role in the ALA since 1988 (c.f. Landau et al., 1988). Words are “in-the-task” contextual cues (like legs and eyes) that dynamically organize attention weights. Words also are important as elements in the system of learned associations (see, especially, Colunga & Smith, 2005; Yoshida & Smith, 2003a, 2003b, 2003c, 2005). For these reasons, ALA predicts and is supported by our studies of cross-linguistic differences. Thus, Booth et al.’s (2005) demonstration that linguistic cues affect children’s noun extensions is not evidence against the ALA. The ALA is not a bottom-up theory of categories; instead, it is explicitly about how knowledge (learned associations among words, perceptual properties, and category organization) exerts top-down control on perception through attention and, in so doing, constrains learning in the moment.

Booth et al. (2005) also misrepresent our work, and that of others, in their assertion that there have been “failures to replicate” (Booth et al., 2005, p. 493) a link between vocabulary growth and the shape bias. The phrase “failure to replicate” implies that investigators used our same experimental procedures to investigate the same empirical question and obtained different results. This is not the case. Diesendruck and Bloom (2003), which was cited as one such “nonreplication,” did not test the relation between the shape bias and early vocabulary growth but instead showed that 2-
and 3-year-olds’ attention to shape could be manipulated by aspects of the task context (which they interpret as indicating a conceptual, rather than perceptual, basis to the shape bias but which we have interpreted in terms of a system of learned associations; see Samuelson & Smith, 2000b). Graham and Poulin-Dubois (1999), which was cited as the second “nonreplication,” used a preferential looking paradigm to examine 16- to 20-month-olds’ attention to shape versus attention to color for pictures of animals versus artifacts. They found that children with both small and large vocabularies systematically attended to shape in both cases (not attending to color more for the animals than for the artifacts), a result consistent with our view of the greater statistical relevance of shape and the extremely limited statistical relevance of color to noun categories (see Samuelson & Smith, 1999). Booth et al.’s (2005) assertion of “nonreplicability” with regard to these papers is disingenuous. Neither study attempted, nor was motivated as, a direct replication of any of our demonstrations of a link between attention to shape and early vocabulary growth, demonstrations that have used longitudinal, correlational, and training studies with both typically developing children and children with language delay (e.g., Gershoff-Stowe & Smith, 2004; Jones, 2003; Jones & Smith, 2005; Smith, 1995; Smith et al., 2002). Our data, along with those of Diesendruck and Bloom (2003) and Graham and Poulin-Dubois (1999), require the unified explanation that the ALA offers, not dismissal. “Failure to replicate” is a misleading choice of words that does a disservice to scientific progress.

What Is at Stake?

The present debate on the shape bias is so unclear that proponents of both sides can conduct nearly identical experiments and see the same results as being for and against the ALA (compare Booth & Waxman, 2002, with Yoshida & Smith, 2003b; Cimpian & Markman, 2005, with Yoshida & Smith, 2003a; and Diesendruck & Bloom, 2003, to Samuelson & Smith, 2000b)! Just what is at issue here?

Traditional theories of cognition partition mental life into three mutually exclusive parts: sense, think, and act (e.g., Fodor, 1983; Keil, 1994; Newell & Simon, 1972). Cognition under this framework is predominantly about the “think” part, which mediates between perceiving and acting. In this view, the knowledge that mediates sensation and action generally is believed to be propositional and symbolic and thus profoundly different from the processes that underlie perceiving, remembering, attending, and acting. Further, the central theoretical goal in this tradition is specification of knowledge structures. For this reason, sense-think-act theories of cognition often are amechanistic and are not focused on the processes that underlie real-time performance. It is under this theoretical set of assumptions that the distinctions between competence versus performance, conception versus perception, and representation versus process make sense (see Thelen & Smith, 1994).

The ALA is derived from entirely different assumptions. This newer idea, one with growing support from computational, behavioral, and cognitive neuroscience approaches to cognition, is that knowledge is embedded in, distributed across, and inseparable from the real-time processes of perceiving, remembering, attending, and acting (see Samuelson & Smith, 2000a). Cognition just is a set of internal processes bound to one another and to the world through perception and action in real time with no fixed and segregated representation of anything. Thus, the ALA does not favor perception over conception, performance over competence, or process over representation because it makes no such distinctions at all. We also make no claims about children’s “understanding” of categories as a unitary process. Understanding is a momentary event embedded in the real-time processes of perceiving, attending, and remembering. Under this newer framework, the explanatory power formerly attributed to concepts as unitary representational entities dissolves, as cognition is emergent in the task and the soft assembly of many processes, each bound to one another and to the here and now of sensorimotor input (see Barsalou, 1993, 2003; Jones & Smith, 1993; O’Regan & Noe, 2001; Port & van Gelder, 1995; Samuelson & Smith, 2000a; Smith, 2005; Smith & Gasser, 2005; Thelen & Smith, 1994).

What is at stake in these debates about the shape bias is the foundational nature of cognition itself. In this larger theoretical battle, our goal is to show how knowledge is embedded in processes such as attention rather than in fixed representational structures such as concepts. Attention is a potentially powerful repository of knowledge because it is a dynamic process shifting continually with changes in the immediate input and because it is informed by a system of learned associations that reflect the statistical regularities of the world, of language, and of the social structure of discourse. Attention is a particularly powerful force on development because it selects and constrains what is learned in the moment, in the real time in which learning and development genuinely take place.

References


Gathercole, V. C. M. (1997). The linguistic mass/count distinction as an


Received September 6, 2005
Revision received February 5, 2006
Accepted February 7, 2006