



Cognitive Science 34 (2010) 1287–1314

Copyright © 2010 Cognitive Science Society, Inc. All rights reserved.

ISSN: 0364-0213 print / 1551-6709 online

DOI: 10.1111/j.1551-6709.2010.01130.x

Knowledge as Process: Contextually Cued Attention and Early Word Learning

Linda B. Smith,^a Eliana Colunga,^b Hanako Yoshida^c

^a*Department of Psychological and Brain Sciences, Indiana University*

^b*Department of Psychology and Neuroscience, University of Colorado*

^c*Department of Psychology, University of Houston*

Received 24 February 2009; received in revised form 24 June 2010; accepted 24 June 2010

Abstract

Learning depends on attention. The processes that cue attention in the moment dynamically integrate learned regularities and immediate contextual cues. This paper reviews the extensive literature on cued attention and attentional learning in the adult literature and proposes that these fundamental processes are likely significant mechanisms of change in cognitive development. The value of this idea is illustrated using phenomena in children's novel word learning.

Keywords: Cued attention; Development; Noun learning

1. Introduction

In her introduction to the 1990 special issue of *Cognitive Science*, Rochel Gelman asked, “How is it that our young attend to inputs that will support the development of concepts they share with their elders?” Gelman posed the question in terms of attention, but the answers offered in that volume were not about attention. Instead, they were about innate knowledge structures, so-called first or core principles that guide learning in specific knowledge domains. In the years since 1990, inspired in part by the highly influential papers in that special issue, there have been many *demonstrations* of the remarkable knowledge that quite young children bring to bear on learning. But there have been very modest advances in understanding the *processes and mechanisms* through which that knowledge is realized and applied to aid learning. Accordingly, in this paper we return to Gelman's original question: How do children select the right information for learning?

Correspondence should be sent to Linda B. Smith, Department of Psychological and Brain Sciences, Indiana University, 1101 East 10th Street, Bloomington, IN 47405. E-mail: smith4@indiana.edu

We consider this question in the context of how children learn words, one of the topics also of central interest in the original special issue. We begin by reviewing well-documented and general mechanisms of attentional learning from the adult literature from the perspective of their relevance as mechanisms of cognitive development. We then ask—mostly without direct empirical evidence on the answer—what role these known mechanisms might play in early word learning. Our main goal is to encourage researchers to pursue these mechanisms as significant contributors to word learning and to cognitive development more generally. This—taking well-documented mechanisms from one area of research and asking whether they might apply in another, moving the field toward a more unified understanding of what might seem at first unrelated domains—should not be contentious. However, the literature on early word learning is highly contentious (Booth & Waxman, 2002; Cimpian & Markman, 2005; Smith & Samuelson, 2006) and not in a productive way, as the explanations pitted against each other are not empirically resolvable in a straightforward way.

One word-learning principle, mutual exclusivity, discussed in the original special issue may help illuminate the problem. The phenomenon, which is not at issue, is this: Given a known thing with a known name (e.g., a cup) and a novel thing and told a novel name (“Where is the rif?”) children take the novel word to refer to the novel object and not the known one. Markman (1990, p. 66) explained this behavior as follows: “children constrain word meaning by assuming at first that words are mutually exclusive—that each object can have one and only one label.” This description summarizes the phenomenon in terms of the macro-level construct of “assumptions.” This construct, of course, may be unpacked into a number of micro-level processes, including, as we will propose here, cue competitions. An account in terms of cue competitions and attention is not in opposition to an account in terms of assumptions because the two kinds of explanations are at fundamentally different levels of analysis that answer different questions. The mutual exclusivity assumption as proposed by Markman is a statement about an operating characteristic of the child’s cognitive system that facilitates word learning. A proposal about cue competitions is a proposal about the more micro-level mechanisms that may give rise to that operating characteristic. An analogy helps: One possible account of why someone just ate a cookie is that they are hungry; another possible account is that they ate the cookie because of low levels of leptin and the release of ghrelin. The second account might be wrong, and the first might well be right; but the second is not in opposition to the first in any sensible way. Moreover, a choice between the two cannot be made according to which one better accounts for the macro-level behavior of eating the cookie. Instead, the relevance of leptin and ghrelin to cookie eating must be decided in terms of micro-level processes about which the macro-level construct of hunger makes no predictions.

In what follows, we consider Gelman’s original question of how children might know to attend to the right properties for learning by considering contemporary evidence on attentional mechanisms and then asking whether these mechanisms might play a role in enabling children to attend to the right information for learning words. The evidence suggestive of a role for cued attention in word learning is primarily at a macro level. Accordingly, we next discuss the kind of micro-level studies needed to pursue these proposed mechanisms. We do not consider macro-level explanations of children’s early word learning as competing

hypotheses to the proposals about cued attention. However, we conclude with a reconsideration of the contention in the early word learning and what is (and is not) at stake.

2. Cued attention

Attention is a construct that is so widely used in psychological theory that William James (1890) lamented “everyone knows what attention is” with the subtext that everyone may know but no one agrees. Within developmental psychology, “attention” is currently studied with respect to several different (but potentially deeply related, see Posner & Rothbart, 2007) phenomena, including sustained attention (e.g., Miller, Ables, King, & West, 2009; Richards, 2005, 2008), attentional switching and disengagement (e.g., Blaga & Colombo, 2006; Posner, Rothbart, Thomas-Thrapp, & Gerardi, 1998; Richards, 2008), executive control (e.g., Chatham, Frank, & Munakata, 2009; Diamond, 2006; Hanania & Smith, in press), and joint attention among social partners (e.g., Grossmann & Farroni, 2009; Hirotani, Stets, Striano, & Friederici, 2009). There are very few developmental studies specifically concerned with contextually cued attention (e.g., Goldberg, Maurer, & Lewis, 2001; Smith & Chatterjee, 2008; Wu & Kirkham, in press). Although we will briefly consider how cued-attention might be related to other forms of attention in children at conclusion, we focus on contextually cued attention in this paper precisely because so little is known about its development despite its apparent ubiquity in sensory, perceptual, and cognitive processing in adults. Briefly, the well-documented fact is this: Cues that have been probabilistically associated with some stimulus in the past enhance detection, processing, and learning about that stimulus (e.g., Brady & Chun, 2007; Chun & Jiang, 1998). In this section, we briefly summarize the very broad literature that supports these conclusions, noting how these processes capture regularities in the input, protect past learning, and guide future learning.

2.1. Predictive cues

The attentional consequences of learned associations between cues and the stimuli they predict has been well known since the work of Mackintosh (1975) and Rescorla and Wagner (1972). Originally in the context of classical conditioning but more recently also in the broader domain of associative learning (see Chapman & Robbins, 1990; Kruschke & Blair, 2000; see also, Ramscar, Yarlett, Dye, Denny, & Thorpe, in press), these studies present learners with cues that predict specific outcomes. The results show that what is learned about the relation between those cues and outcomes depends on the cues present in the task, their relative salience, *and the learner’s history of experiences with those cues in predicting outcomes*. Critically, these cued-attention effects are not about single cues associated with single attentional outcomes but rather are about the consortium of cues present in the moment and all of their predictive histories.

Three illustrative phenomena are overshadowing, blocking, and latent inhibition (Kamin, 1968; Lubow & Moore, 1959; Mackintosh, 1975; Rescorla & Wagner, 1972). The relevant task structures are shown in Table 1. Overshadowing refers to the situation in which two

Table 1
Cue interactions in associative learning

Cue Interaction	First Learning Phase	Second Learning Phase	Learning Outcome
Overshadowing	AB → X		A → X
Blocking	A → X	AB → X	A → X
Latent inhibition	B →	AB → X	A → X
Highlighting	AB → X	AC → Y	C → Y
Mutual exclusivity	A → X	AB → Y	B → Y

Note. The letters A, B, C indicate cues and the letters X and Y indicate predicted outcome for the first phase of learning, and the subsequent second phase of learning. What is learned after the second phase is indicated in the third column. Bold letters indicate more salient cues (either through learning or intrinsic salience).

cues are presented together and jointly predict some outcome (e.g., in a category learning study, two symptoms might predict some disease, or in classical conditioning, a tone and a light might predict shock). The strength of association of each cue to the outcome depends on its *relative* salience. However, it is not simply that the most salient cue is learned better; rather, salience differences are exaggerated in that the more salient cue may “overshadow” the less salient cue with little or no learning at all about the less salient cue (Grossberg, 1982; Kamin, 1969; Kruschke, 2001; Mackintosh, 1976; Rescorla & Wagner, 1972). Blocking refers to the case in which the greater salience of one cue is not due to intrinsic salience but is due, instead, to past learning. If some cue regularly predicts some outcome and then, *subsequently*, a second cue is made redundant with the first and so also predicts that outcome, there is little or no learning about the second cue: Learning is blocked by the first predictive cue (Bott, Hoffman & Murphy, 2007; Kamin, 1968; Kruschke & Blair, 2000; Shanks, 1985), as if the first cue were more “salient” and thus overshadowing in this predictive context. Latent inhibition, like blocking, also makes the point that salience is a product of learning and predictive strength. But here the phenomenon is learned irrelevance: If a cue is first varied independently of the outcome so that it is not predictive at all but then it is made perfectly predictive, it will be hard to learn and “overshadowed” by other cues (e.g., Lubow, 1997; Lubow & Kaplan, 1997; Mackintosh & Turner, 1971).

These phenomena make three key points about the cues in cued attention: (a) cue strength is determined by predictive power, not mere co-occurrence; (b) cue strength depends on the ordered history of predictability, not batch statistics; and (c) individual cues interact, such that one cannot simply predict whether some cue will be learned by considering it alone, one must instead know its history and the history of the other cues in the learning environment. These phenomena are evident in many adult statistical learning tasks (e.g., Cheng & Holyoak, 1995; Kruschke, 2001; Kruschke & Blair, 2000; Kruschke & Johansen, 1999; Ramscar et al., in press) and as Ellis (2006) concluded, their ubiquity implies that human learning about predictive relations is bounded by basic mechanisms of cued attention (see Yu & Smith, in press; Yoshida, unpublished data).

Highlighting is a higher level phenomenon that may be understood as a product of blocking and overshadowing and it is a particularly robust phenomenon in adult associative learning (Kruschke, 1996, 2005; Ramscar et al., in press). It also demonstrates how cued

attention both protects past learning and guides new learning. Highlighting emerges when adults first learn one set of predictive cues to task-relevant information and then later are exposed to an overlapping set of *new and old* predictive cues that predict the relevance of *different* information. The task structure that leads to highlighting is also provided in Table 1. Learners are first exposed to a conjunctive cue (A + B) that predicts an outcome (X), and then are presented with a new conjunctive cue that contains one old component (A) plus a new one (C) that predicts a new outcome (Y). The key result concerns learning during the second phase: Learners associate the new cue with the new outcome more than they associate the old cue with the new outcome. For example, if the learner is first taught that RED and SQUARE predict category X and then is taught that RED and CIRCLE predict category Y, the learner does not learn about the relation of RED and category Y, but rather appears to selectively attend only to the novel cue, CIRCLE, and to learn that CIRCLE predicts category Y. In brief, novel cues are associated with novel outcomes. By one explanation, this derives from the rapid (and automatically driven) shift of attention away from the previously learned cue (RED and SQUARE) in the context of the new outcome, so that the new cue (CIRCLE) becomes attentionally highlighted and strongly associated with the new outcome (e.g., Kruschke, 1996, 2005). This attention-shifting account has also been supported by eye-tracking results (Kruschke, Kappenman, & Hetrick, 2005).

Highlighting, as well as blocking and overshadowing may be understood in terms of competitions among predictive cues, as if cues fight for associations with outcomes, so that once an association is established it protects itself by inhibiting the formation of new associations to the same predicted outcome. There are many reasons to think that these kinds of cue competitions should play a role in cognitive development, including the robustness of the phenomena in adults across a variety of task contexts. There is also one phenomenon in children's learning that shares an at least surface similarity to cue competition effects in associative learning: mutual exclusivity (Markman, 1989; see also Halberda, 2006; Hollich et al., 2000).

As illustrated in Table 1, the task structure that yields mutual exclusivity looks very much like that of blocking or highlighting: A first-learned association (the word "cup" to the referent cup) is followed by a subsequent learning task with a new cue (the novel word) and outcome (the novel object). Consistent with this proposal, several models (e.g., Mayor & Plunkett, 2010; Regier, 2005) have shown how mutual exclusivity might emerge in cue interactions in associative learning. Also consistent with this proposal are studies tracking moment-to-moment eye gaze direction in infants (Halberda, 2009); the attention shifting by infants in these studies resembles those of adults in highlighting experiments (Kruschke et al., 2005).

2.2. Lots of cues

Blocking and highlighting emerge in experimental studies in which researchers manipulate at most two to three cues and outcomes. The world, however, presents many probabilistically associated cues and outcomes yielding potentially complex patterns of predictability, with cues predicting other cues as well as potentially multiple outcomes. Importantly, such a

large “data set” of associations are also likely to have considerable latent structure, higher order regularities that might support deep and meaningful generalizations that go beyond specific cues and outcomes. Several connectionist models have sought to understand the structure of these higher order regularities (see, e.g., Colunga & Smith, 2005; Colunga, Smith, & Gasser, 2009; Kruschke, 1992, 2001; McClelland & Rogers, 2003). In one study, Colunga and Smith (2005) examined how the statistical regularities among perceptual properties of instances of basic level noun categories might create higher order partitions (object, substance) and also direct attention to relevant properties for categorizing artifacts (shape) and substances (material). To this end, they fed a connectionist net perceptual features (e.g., solidity, material, color, shape) that adults said characterized 300 noun categories commonly known by 2-year-olds. From this input, the network acquired generalized cued-attention effects that worked even when presented with novel entities; specifically, in making decisions about novel categories, the network weighted shape more in the context of solidity and weighted material in the context of nonsolidity. This result tells us that the corpus of early learned nouns and the correlated properties of the objects to which those nouns refer contain useable cue-outcome regularities of the kind that could train attention.

Several theoretical and empirical analyses suggest that systems of associations often present a form of *coherent covariation* (Rogers & McClelland, 2004) across cues and outcomes. The importance of coherent covariation to human learning has been demonstrated in experimental tasks showing that adults (and infants, Younger & Cohen, 1983) are more likely to attend to and learn about features that co-vary than those that merely co-occur (e.g., Kruschke & Blair, 2000; Medin & Schaffer, 1978; Medin, Altom, Edelson, & Freko, 1982). Other theorists have also pointed to what they call the *systematicity* (Billman & Heit, 1989) of associations; cues that probabilistically predict an outcome also often predict each other, and the systematicity across associations among cues, among outcomes, and among cues and outcomes matters in learning (e.g., Billman & Knutson, 1996; Goldstone, 1998).

Yoshida and Smith (2003b, 2005) (see also, Sloutsky & Fisher, 2008) proposed a cued-attention framework for thinking about the interactive effects of redundant cues that builds on the idea of interactive activation (see Billman & Knutson, 1996; Goldstone, 1998; Medin et al., 1982; O’Reilly, 2001). The proposal is illustrated in Fig. 1 using the simple case of

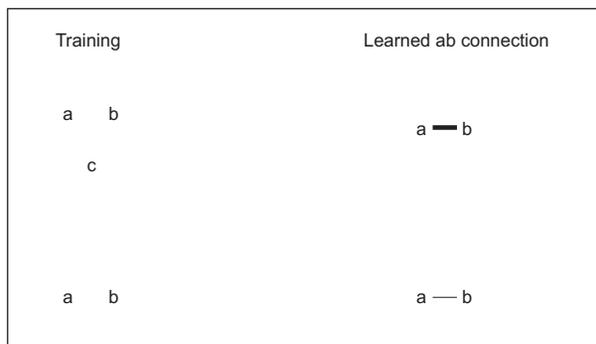


Fig. 1. Redundant correlations in the learning environment lead to stronger associations.

just three correlated cues: The correlation between cues **a** and **b** are learned either in the context of a third redundant cue, **c**, or without that third correlated cue. Experimental studies show that the learned connection between **a** and **b** is stronger if acquired in the context of **c**, which correlates with both **a** and **b**, than without that redundant correlation (Billman & Knutson, 1996; Yoshida & Smith, 2005). Moreover, the stronger associative link between **a** and **b** remains even when the redundant cue, **c**, is removed. Of course, in real-world learning, there may be many more than three correlated cues and much more complex patterns of correlation. Through mutually reinforcing correlations such as these, a system of many correlated cues may lead to what have sometimes been called “gang effects”: Highly interconnected and dense patterns of associative links that give rise to patterns of internal activation that, as a result, are not tightly dependent on any one cue (e.g., Billman & Knutson, 1996; Goldstone, 1998; O’Reilly, 2001; Yoshida & Smith, 2003a,b). Instead, the interactive activation among co-varying cues can lead to snowball effects in which the joint activation of cues is stronger than the sum of the individually contributing associations.

One developmental result raised by Gelman (1990) in the original special issue, “illusory projections,” may be understood in terms of such gang effects (see Rogers & McClelland, 2004 for a more detailed account). Gelman reported that preschool children project properties onto an instance of a category that are not perceptually there; children say that they *saw* feet on an eyed yet ball-like and footless thing if they were told that the thing could move on its own. Fig. 2 provides an example of how this might work within the cued-attention framework. On the left is a hypothetical description of overlapping correlations among a set of perceptual properties: Things with eyes tend also to have feet, to move on their own, and so forth. Because of the overlapping correlations among all these properties, each will serve as a context cue that predicts and “primes” attention to the others (see Yoshida & Smith, 2003b) thereby potentially causing an individual to more readily detect or “see” a property in an ambiguous stimulus. Thus, on the right of Fig. 2 is a hypothetical description of the cluster of cues input to the system—eyed, furry, moving on its own, but footless. By hypothesis, fur, eyes, and movement will all increase attention to each other and, given the ambiguous stimulus, lead to the perception of feet.

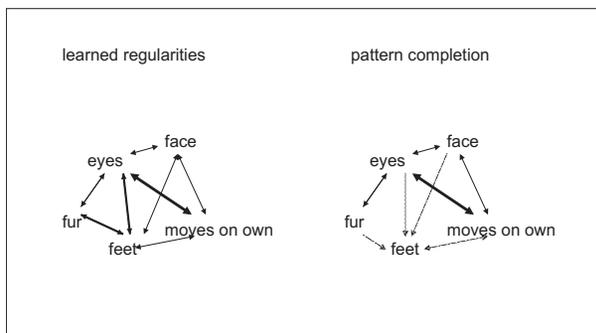


Fig. 2. An illustration of how redundant and overlapping correlations can lead to pattern completion and illusory projections.

2.3. *Enhanced processing of predicted outcomes*

Cues predict outcomes, and in the present use of the term, an outcome is any stimulus event that is predicted (it could be another cue or the specifically task-relevant target information). Considerable and growing evidence suggests that cues that predict outcomes also enhance detection and processing of the predicted event. The relevant experiments often involve search or detection tasks in which participants are asked to detect some target—a shape or color—typically in a crowded field of competing distractors (Chun & Jiang, 1998; Jiang, Olson, & Chun, 2000). These studies have revealed strong effects of repeating arrays. Arrays that have been seen before show enhanced detection of the target, often with just one pre-exposure but increasing with repetition. This phenomenon is usually explained as a form of cued attention in which the array as a whole cues attention to a specific target at a specific location (Biederman, 1972; Boyce, Pollatsek, & Rayner, 1989; Lewicki, Hill, & Czymowska, 1992; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977).

The experimental evidence also shows that in these tasks, contextual cues and their attentional effects emerge without awareness (Chun & Jiang, 1998; Jiang & Chun, 2001, 2003; Jiang & Leung, 2005; Lewicki, Hill, & Czymowska, 1997; Lewicki et al., 1992; Olson & Chun, 2002; Shanks, Channon, Wilkinson, & Curran, 2006) and result both in rapid shifting of attention to a location (e.g., Chun and Turke-Browne, 2007; Clohessy, Posner, & Rothbart, 2001; Jiang & Chun, 2001; Summerfield et al., 2006) and also to the enhanced processing of particular stimulus features (Bichot & Rossi, 2005; Kruschke, 1996; Maunsell & Treue, 2006; Rossi & Paradiso, 1995). The growing neuroscience evidence on cued attention also indicates that contextually cued enhancements of stimulus processing are pervasive across early sensory processing and higher level perceptual and cognitive systems (Beck & Kastner, 2009; Gilbert, Ito, Kapadia, & Westheimer, 2000; Pessoa, Kastner, & Ungerleider, 2003). For example, neurophysiological studies have shown that cue-target associations enhance baseline firing rates (Chelazzi et al., 1993, 1998; Kastner et al., 1999) at early and middle levels in the visual system and alter neuronal tuning of target properties (Spitzer et al., 1988; Treue & Martinez Trujillo, 1999; Williford & Maunsell, 2006; Yeshurun & Carrasco, 1998). Context cued enhancements of processing have also been shown to play a role in decision making and in integrating sensory information across systems (Bichot et al., 1996; Boynton, 2009; Gold & Shadlen, 2007; Reynolds & Heeger, 2009).

Contemporary theories of these effects, often based on analyses of neural patterns of excitation, share a great deal with more classical explanations of cued attention (e.g., Mackintosh, 1975; Rescorla & Wagner, 1972) that emphasize prediction (or preactivation, e.g., Summerfield et al., 2006) and competition (e.g., Desimone & Duncan, 1995; Duncan, 1996). For example, the biased-competition theory of selective attention (see Beck & Kastner, 2009; Desimone & Duncan, 1995; Duncan, 1996) begins with the starting assumption that competition characterizes representational processes at the sensory, motor, cortical, and subcortical levels. In general, activation of a representation of some property, event, or object is at the expense of other complementary representations. Selection, or attention, then, occurs by biasing (e.g., priming) some representations in favor of others or by inhibiting competing repetitions. The presence of contextual cues previously associated with some

representation (at any of these levels) is thus thought to bias the competition. Multiple activated cues compete, and thus also interfere with each other, with stronger cues inhibiting weaker ones in a manner common to lateral inhibition models (Ludwig, Gilchrist, & McSorley, 2005; Walley & Weiden, 1973; de Zubicaray & McMahon, 2009). On these grounds, some have suggested that attention and selection are fundamentally a form of lateral inhibition in which the degree of activation of one representation inhibits that of nearby competitors (see Beck & Kastner, 2009; Duncan, 1996).

Here, then, is what we know about cued attention: Adults readily and unconsciously learn cues that predict other sensory events (outcomes) that are relevant in tasks. These predictive cues interact—both by supporting activation of the predicted event when they are correlated and also through competition that depends in fine-grained ways on the relative strengths of these cues and their history of prediction. These interactions are such that (a) early learned predictive relations tend to be preserved; (b) attention is systematically shifted to novel cues in the context of novel outcomes; and (c) the coherent covariation in large systems of cues and outcomes can capture latent structure that organizes attention in meaningful ways. Finally, predicted sensory events are processed faster and tuned more sharply than unpredicted events. All this suggests that cued attention is a basic mechanism that is likely to play a contributing role in many knowledge domains.

3. Cued attention as developmental process

Cued attention is also a mechanism of change and one that seems capable of driving considerable change in the cognitive systems. This is because cued attention is a single mechanism that aggregates knowledge, that is a repository of knowledge, and that guides learning, thereby driving the acquisition of new knowledge. In brief, attentional learning is a self-organizing process that builds on itself, becoming more directed, more knowledge driven, and potentially more domain specific as a consequence of its own activity. Fig. 3 builds on a prior proposal by Colunga and Smith (2008) about how cued attention gathers, integrates, and applies information over nested time scales. The three separate boxes on the left illustrate attentional processes at three time scales. The large box represents long-term associations among the many cues and outcomes in the learning environment. The middle box indicates the task context and the in-task dynamics of prediction and preactivation of cues and outcomes (see Samuelson & Smith, 2000a, 2000b; Samuelson, Schutte, & Horst, in press). Finally, there are the processes of interaction and competition driven by the current task and stimuli. As illustrated on the right side of the figure, these are nested processes: In-the-moment attention depends on task context as *modulated* by long-term associations; in-the-moment attention also adds to those long-term associations.

These nested interactions mean that attention can be biased in different ways in different contexts. Context cues that co-occur with (and define) specific tasks will come with repeated experience to shift attention to the task-relevant information. Fig. 4 represents this idea in terms of a series of context-sensitive salience maps, with the relative salience of regions in the perceptual field indicated by the relative darkness of the locations. The idea is this:

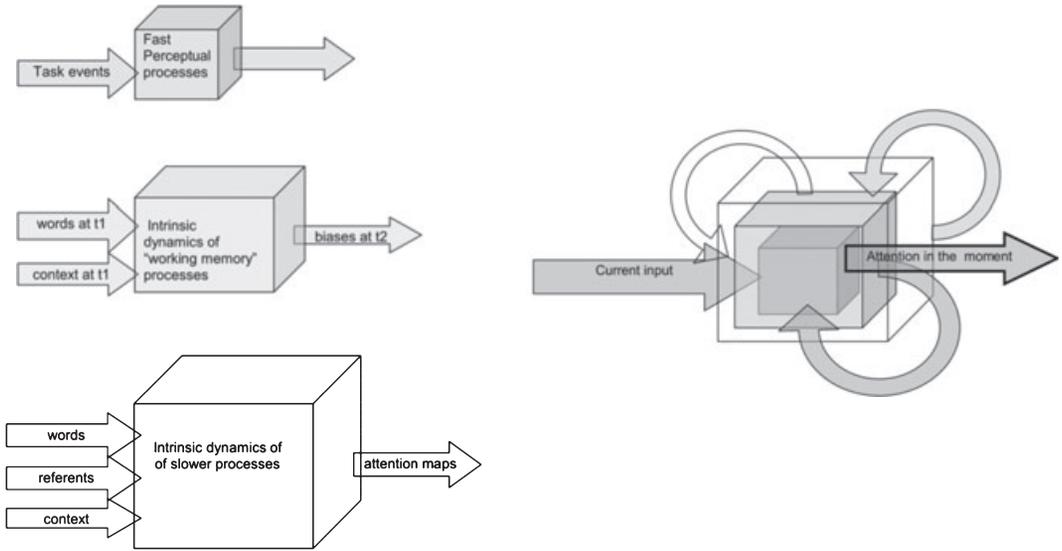


Fig. 3. The nested time scales of attentional learning. The three separate boxes on the left illustrate attentional processes at three time scales. The large box represents long-term associations among the many cues and outcomes in the learning environment. The middle box indicates the task context and the in-task dynamics of prediction and preactivation of cues and outcomes. The small box indicates processes of interaction and competition to a momentary stimulus. As illustrated on the right side of the figure, these are nested processes: In-the-moment attention depends on task context as *modulated by* long-term associations; in-the-moment attention also adds to those long-term associations.

Because the history of associated cues increases and decreases the relative activation of features and task targets in the perceptual field, the salience map *will* change with changes in contextual cues. This means potentially dramatic shifts in the detection, selection, and processing of stimulus events in different domains of expertise—word learning, quantity judgments, or spatial reasoning. These domains have associated contextual cues that—given sufficient experience—may structure the salience maps in consequentially different ways. Being smart in a domain may reflect (in part) predicting what information is relevant in that domain. For learners who have sufficient experiences in different domains, attention will nimbly dance about from context to context, enabling those learners to attend to just the right sort of information for that domain. In sum, what we know about cued attention and attentional learning suggests that these mechanisms will play a role across all domains of cognition and their self-changing nature will create domain-specific competencies. We consider next one domain in which these mechanisms may be at work.

4. Children's novel word generalizations

Many well-controlled studies show that young children need to hear only a *single* object name to systematically generalize that name to new instances in ways that seem right to

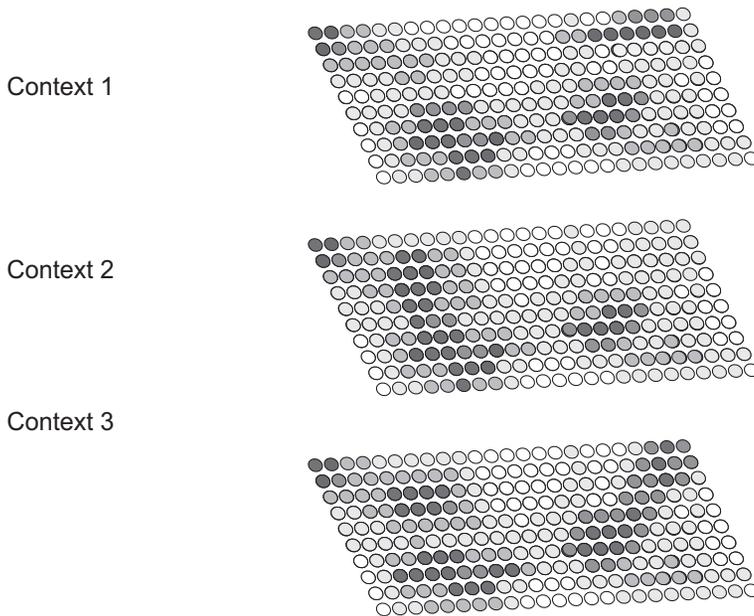


Fig. 4. Contextually changing salience maps. Relative salience is indicated by the relative darkness of the locations. By hypothesis, these change with changes in context cues that predict sensory events in the field.

adults (e.g., Golinkoff, Mervis, & Hirsh-Pasek, 1994; Markman, 1989; Smith, 1995; Waxman & Markow, 1995). Moreover, children generalize names for different kinds of things by different kinds of similarities, shifting attention to the right properties for each kind of category. Thus, for the task of learning common nouns, young children have solved Gelman's problem: They know what properties to attend to so as to form categories "shared with their elders."

In the most common form of the novel word generalization task, the child is shown a single novel entity, told its name (e.g., *This is the toma*) and then asked what other things have the same name (e.g., *Where is the toma here?*) Many experiments have examined three kinds of entities (examples are shown in Fig. 5) and found three different patterns of generalization. Given objects with features typical of animates (e.g., eyes or legs), children extend the name narrowly to things that are similar in multiple properties. Given a solid inanimate artifact-like thing, children extend the name broadly to all things that match in shape. Given a nonsolid substance, children extend the name by material. These are highly reliable and replicable results—obtained by many researchers—and in their broad outline characteristic of children learning a variety of languages (e.g., Booth & Waxman, 2002; Gathercole & Min, 1997; Imai & Gentner, 1997; Jones & Smith, 2002; Jones, Smith, & Landau, 1991; Kobayashi, 1998; Landau, Smith, & Jones, 1988, 1998; Markman, 1989; Soja, Carey, & Spelke, 1991; Yoshida & Smith, 2001; see also Gelman & Coley, 1991; Keil, 1994).

Cued attention may play a role in these generalizations in the following way: Artifacts, animals, and substances present different features (angularity, eyes, nonsolidity) and these

Proportions of Name Generalizations

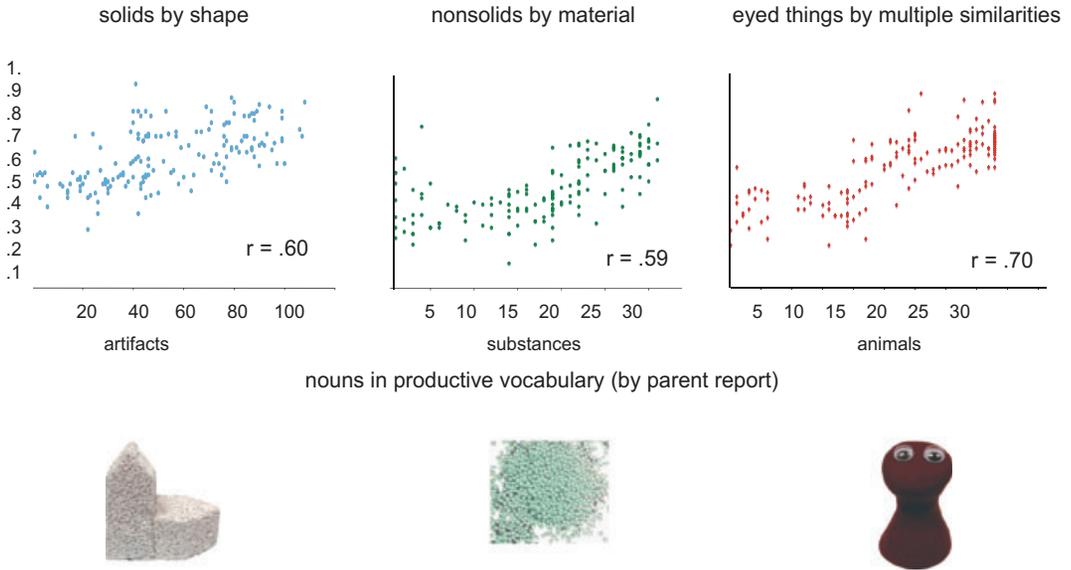


Fig. 5. Scatterplots of individual children's performances in the Novel Noun Generalization task as a function of the number of different kinds of nouns in their productive vocabularies. Individual data are from: Smith et al., 1992; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002; Yoshida & Smith, 2003b; Jones & Smith, 1998, 2002; and Jones & Smith, unpublished data.

features co-occur with different words (quantifiers, verbs, adjectives, etc.). The features and co-occurring words are thus potential context cues that could shift attention in systematic ways to the relevant properties for the particular kind of category—to multiple properties for animals to shape for artifacts, and to material for substances. And, indeed, the literature is filled with experimental demonstrations of these effects (Booth & Waxman, 2002; Colunga, 2006; Colunga & Smith, 2004; Gathercole, Cramer, Somerville, & Jansen op de Haar, 1995; Jones & Smith, 1998; McPherson, 1991; Samuelson, Horst, Schutte, & Dobbertin, 2008; Soja, 1994; Ward, Becker, Hass, & Vela, 1991; Yoshida & Smith, 2005; Yoshida, Swanson, Drake, & Gudel, 2001).

These perceptual- and linguistic-context effects on children's novel noun generalizations also increase with age and language learning, just as they should if children are learning relevant contextual cues (Jones et al., 1991; Landau et al., 1988; Samuelson, 2002; Samuelson & Smith, 1999, 2000a; Smith, 1995; Soja et al., 1991). Fig. 5 presents a summary of the developmental pattern. The figure shows scatterplots of individual children's novel noun generalizations for solid artifactual things, things with eyes, and nonsolid things from a variety of different experiments (with many different unique stimuli and task structures) conducted in our laboratories over the years. Each individual child's generalizations by the category-relevant property (shape for solid things, multiple similarities for eyed things, material for nonsolid things) is shown as a function of the number of artifact, animal, and

substance names in the individual child's vocabulary. The figures show that the systematicity of these novel noun generalizations increases with early noun learning. Thus, in the broad view, the developmental pattern of children's novel noun generalizations fits what might be expected if the underlying mechanism was cued attention: Children learn cues for task-relevant properties and after sufficient experience, those cues come to shift attention in task-appropriate ways.

5. Cross-linguistic differences

By hypothesis, the learning environment presents clusters of perceptual cues (e.g., eyes, legs, mouths, body shapes in the case of animates) and clusters of linguistic cues (e.g., the words "wants," "is happy," "mad," and "hungry") that are associated with each other and that predict the relevant similarities for categorizing different kinds. If perceptual and linguistic cues are both part of the same attentional cuing system they should interact, and given coherent covariation, should reinforce each other. The "natural" experiment that provides evidence is the comparison of children learning different languages.

5.1. Systematicity

The novel word generalizations of children learning English and Japanese have been examined in a number of studies (Imai & Gentner, 1997; Yoshida & Smith, 2001, 2003a, 2005). Both languages provide many linguistic cues that correlate with artifacts (and attention to shape), animals (and attention to multiple similarities), and substances (and attention to material). However, English arguably provides more *systematic* and coherently co-varying cues distinguishing objects and substances, whereas Japanese arguably provides more coherently co-varying cues distinguishing animates and inanimates (see Yoshida & Smith, 2003b, for a discussion).

In particular, the count-mass distinction in English partitions all nouns into discrete countable entities (objects and animals) or masses (substances) and the various linguistic markers of this distinction (determiners, plural) are strongly correlated with the perceptual cues (and particularly solidity) that predict categorization by shape. Moreover, the predictive relations between linguistic cues and perceptual cues are particularly strong in the 300 English nouns that children normatively learn by 2½ years (see Colunga & Smith, 2005; Samuelson & Smith, 1999; Smith, Colunga, & Yoshida, 2003). By hypothesis, these linguistic cues should augment attention to shape for solid artifactual things and to material for nonsolid substances. Because these linguistic cues lump object and animal categories together, they might also weaken the predictability of the distinction between artifact categories as shape-based categories and animal categories as organized by multiple properties.

Japanese, unlike English, makes no systematic distinction between count and mass nouns (and has no English-like plural). Therefore, there is less redundancy in the cue-category correlations with respect to object and substance categories. However, Japanese offers more systematic cues with respect to animates and inanimates than does English. As just one

example, every time a Japanese speaker refers to the location of an entity, in frames as ubiquitous as *There is a ____*, they must mark the entity as animate or inanimate (*ga koko ni iru* vs. *__ga koko ni aru*, respectively). In this, as well as other ways, Japanese, relative to English, adds extra and systematic cues that correlate with perceptual cues predictive of animal versus nonanimal category organizations.

Under the cued-attention framework, the coherent variation of linguistic and perceptual cues within the two languages should create measurable cross-linguistic differences in the noun generalizations of children in the artificial word-learning task. For children learning both languages, solidity predicts attention to shape, nonsolidity predicts attention to material, and features such as eyes and feet predict attention to multiple similarities. But for children learning English, the solidity–nonsolidity cues co-vary and are supported by linguistic cues. For children learning Japanese, eyed–noneyed cues correlate with pervasive linguistic contrasts. Thus, there should be stronger, earlier, and sharper distinctions in novel noun generalizations for solid versus nonsolids for English-speaking children than for Japanese-speaking children and stronger, earlier, and sharper distinctions between eyed and noneyed stimuli for children learning Japanese. Experimental studies have documented these differences (Imai & Gentner, 1997; Yoshida & Smith, 2003b).

5.2. *Gang effects*

Yoshida and Smith (2005) provided an experimental test of the cued-attention account of these cross-linguistic differences in a 4-week training experiment that taught monolingual Japanese children redundant linguistic cues analogous to count-mass cues in English. The experiment used a 2×2 design: Linguistic cues versus no linguistic cues correlated with category organization and solidity during training, and the presence or absence of those linguistic cues at test. The key results are these: Children who were trained with correlated linguistic and perceptual cues outperformed those who were not so trained, attending to the shape of solid things and the material of nonsolid thing and did so even when they were tested with totally novel entities *and even when the trained linguistic cue was not present at test*. That is, learning the links between solidity and nonsolidity and shape and material *in the context of redundant linguistic cues* made the perceptual cue–outcome associations stronger.

5.3. *Illusory projections*

In a related study, Yoshida and Smith (2003a) presented English and Japanese children with novel entities like that in Fig. 6. The protrusions are ambiguous; depending on context, both Japanese- and English-speaking children could see them as legs or as wires. However, when presented in a neutral context, Japanese-speaking children were more likely to see them as leg-like and to generalize a novel name for the exemplar by multiple similarities; English-speaking children were more likely to see them as wires and to generalize a novel name for the exemplar by shape. In brief, Japanese-speaking children showed an enhanced sensitivity to subtle cues vaguely suggestive of animacy, a sensitivity that may be created in



Fig. 6. An ambiguous object: The protrusions may be seen as wires or as limbs.

a history of redundant linguistic-perceptual cues that reinforce attention to each other and in so doing prime the relevant representations for animal-like features.

These cross-linguistic differences indicate that perceptual and linguistic cues interact in a single system that organizes attention in novel word generalization tasks. The findings also suggest that cued attention might be one mechanism behind Whorfian effects in which speakers of different languages are found to be particularly sensitive to and attentive to relations and properties that are lexicalized in their language (see Boroditsky, 2001; Bowerman, 1996; Choi & Bowerman, 1991; Gathercole & Min, 1997; Gentner & Boroditsky, 2001; Levinson, 2003).

6. Competition and prediction

6.1. One attentional pathway

One early study by Smith, Jones, and Landau (1992) using the novel-noun generalization task directly examined cue-competitions by pitting cues against each other. The study examined sentence frames associated with naming objects (*This is a riff*) and frames associated with labeling properties (*This is a riff one*) asking under what contexts the count noun frame would result in increased attention to shape and in what contexts the adjective frame would result in increased attention to color. Children know more nouns and know them earlier than they do adjectives, so the noun frame might be expected to be a stronger attentional cue than the adjective frame. Particularly relevant to possible underlying mechanisms of cued-attention, Smith et al. examined the effects of noun and adjective frames when the colors of the labeled things were dull (olive green), intrinsically salient (glittery silver-gold), or very salient (glittery silver-gold under a spotlight).

The results provide strong support for a competition among cues that interacts with other stimulus-driven (exogenous) pulls on attention. Given dull colors, children generalized novel words presented in both the noun and the adjective frames by shape. Given glittery colors without a spotlight, children generalized novel words in a noun frame by shape but in

an adjective frame by color. Apparently the association between the noun frame and shape is strong enough to overcome competition from glitter, but the pull from glitter helps the weaker adjective frame in guiding attention to the property. Finally, when given glittery colors under a spotlight, children often generalized the word—even in the count noun frame—by color. Children’s performance strongly suggests a winner-take-all competition for attention in which learned cues and other pulls on attention directly interact.

These results also illustrate an important point about attention: It forces a decision and there is but one pathway and one decision. The single pathway means that attention must *integrate* information—from past learning, from task contexts, from immediate input—into a single attentional response. Because attention integrates multiple sources of information in the moment, it will also be flexible and adaptive to the idiosyncracies of specific tasks. The results also underscore the idea that words, whatever else they are, are also cues for attention and play a role in guiding in-the-moment attention.

6.2. Cue strength

Colunga et al. (2009) offered an analysis of how the structure of the Spanish and English count-mass system may yield different cue strengths and as a consequence different patterns of cue interactions. The relevant difference between the two languages can be illustrated by thinking about how speakers can talk about a block of wood. An English speaker talking about a wooden block might say “a block” if he or she is talking about its shape or “some wood” if he or she is talking about its substance. Further, an English speaker cannot say “some block” or “a wood” because “block” is a count noun and “wood” is a mass noun. In contrast, a Spanish speaker, talking about the wooden block, could say “un bloque” (a block) or “una madera” (a wood) when talking about the coherent bounded and shaped thing but would say “algo de madera” (some wood) when talking about the substance irrespective of its shape. That is, count-mass quantifiers are used more flexibly across nouns in Spanish than English to signal different task-relevant construals of the entity (Iannucci, 1952). Note that English has some nouns that work like “madera” in Spanish: “a muffin” predicts the context relevancy of muffin shape, but “some muffin” predicts the context relevancy of muffin substance. But such nouns are not common in English, whereas in Spanish, in principle, all nouns work this way, and in everyday speech many more nouns are used in both count and mass frames than in English (Gathercole & Min, 1997; Gathercole, Thomas, & Evans, 2000; Iannucci 1952). In brief, in Spanish, count-noun syntax is more predictive of attention to shape or material than the noun (*madera*) itself or the solidity of the object. In English, syntax, the noun, and solidity are (not perfectly but) more strongly correlated with each other, and equally predictive of whether the relevant dimension is shape or material.

In light of these differences, Colunga et al. hypothesized that Spanish count-mass syntax should overshadow other cues—solidity, the noun—in children’s novel noun generalizations. In the experiment, monolingual Spanish-speaking children and monolingual English-speaking children were presented with a novel solid entity named with either mass or count nouns, e.g., “a dugo” or “some dugo” and then tested with the two syntactic frames. The syntactic frame had a stronger effect on the Spanish-speaking children’s noun

generalizations than on the English-speaking children's noun generalizations. In other words, Spanish-speaking children learn to attend to mass-count syntax as a way to disambiguate nouns that sometimes refer to shape and sometimes refer to material, and in the context of mass syntax they attend to material even when the items are solid. These results fit the idea that children learn predictive cues and that stronger cues inhibit weaker ones. They also emphasize how individual cues reside in a system of cues and that it is the system of interacting cues that determine attention.

6.3. Protecting past learning

Cue competition privileges old learning as stronger cues inhibit new ones and force attention elsewhere. Yoshida and Hanania (2007) provided an analysis of adjective learning that illustrates how this aspect of cued attention could play a positive role in word learning. The motivating question for their analysis was how, in the context of both a novel adjective and a known noun (e.g., in the context of "a *stoof* elephant"), attention could be properly shifted to a property (such as texture) rather than the shape of the thing. If the word "elephant" is a strong cue for attention to elephant shape, how does the child manage to learn novel adjectives? The answer offered by Yoshida and Hanania is much like the mutual-exclusivity account (Markman, 1989) but is in terms of a cued-attention mechanism rather than children's knowledge about how different kinds of words link to different kinds of meanings.

Their experiments were based on a prior study by Mintz and Gleitman (2002) that showed that the explicit mention of the noun (e.g., the *stoof* elephant) helps children learn the novel adjective relative to conditions in which the noun is not explicitly spoken (e.g., a *stoof* one). Yoshida and Hanania proposed that the role of the noun could be understood as a kind of attentional highlighting via competition. In brief, their cued-attention explanation of the explicit mention of the noun is this: Children usually experience the word *elephant* in the context of elephant-shaped things with typical elephant textures (e.g., *ROUGH*). In these novel adjective-learning experiments, the named object is an elephant-shaped elephant with a totally novel texture (e.g., with tiny holes punched throughout). Thus, in the context of *stoof elephant*, the known cue-outcome (*elephant*-elephant shape) shifts attention to the novel cue—novel outcome (*stoof*-holey texture). In the context of *stoof one*, there is no conjunctive cue containing both a known and a novel cue and thus less competition, which in this case means less directed attention to the novel property. Yoshida and Hanania provided indirect support for this account by showing that the mere conjunction of words (in an unordered list format, e.g., *elephant, stoof, red*) rather than in a sentence format in which *stoof* modifies the known noun (*elephant*) was sufficient for the novel-word to novel-property mapping and by showing that the attentional shift away from shape to the right property, depends on the number of competitive cues (*stoof* in the context *red* and *elephant* leads to stronger effects than *stoof* in the context of elephant alone).

Competitive processes are ubiquitous in the sensory and cognitive system (see Beck & Kastner, 2009). They are at the core of current theories of lexical access and on-line sentence processing (e.g., Bowers, Davis, & Hanley, 2005; Davis & Lupker, 2006). It seems

likely that they play an important role in lexical development as well (Halberda, 2009; Hollich, Hirsh-Pasek, Tucker, & Golinkoff, 2000; Horst, Scott, & Pollard, in press) and, as suggested by Yoshida & Hanania, 2007; Yoshida and Hanania, unpublished data), competitive attentional processes may be particularly important in early development because by protecting already learned associations, they guide learning about novel cues and outcomes, and in this way may leverage little bits of learning into strong forces that effectively speed up learning.

The evidence on children's novel noun generalizations show that context—both linguistic and perceptual—cues children to form categories on the basis of different properties and to select one object over another as the referent of some word. The contextual cueing effects that characterize these early word generalizations share a number of (at least surface) similarities to cued attention—the role of coherent covariation, competition with other pulls on attention, attention shifting to protect past learning. However, none of these early word-learning experiments unambiguously show that these are *attentional* effects. They cannot because the measures of performance are at a macro level—generalizing names to things. To show that mechanisms of attentional learning play a role in these phenomena, we need measures at finer levels of resolution.

7. Going micro

Our main thesis is that cued attention is a potentially powerful developmental mechanism and that early word learning—and particularly the phenomena associated with children's smart novel noun generalizations—would be a fertile domain in which to investigate this idea. However, the granularity of the phenomena studied under cued attention—rapid attentional shifts, rapid detection, enhanced processing, and discrimination—and those studied in most novel word-learning experiments—generalization of an object name to a new instance—are not comparable. What is needed are finer-grained measures of component behaviors—orienting, disengagement, detection, and discrimination—in the context of words and correlated object properties. One might ask, for example, whether a count noun sentence frame primes detection of an odd shape but not an odd texture in a visual search task or whether a count noun sentence frame enhances shape discriminations over texture discriminations.

Recent studies of early word learning using moment-to-moment tracking of eye gaze suggest the new insights to emerge from analyzing attention and word learning at a finer temporal scale. For example, recent studies using this methodology suggest that older (24 month old) and younger (18 month old) word learners may both map words to referents, but that older learners more rapidly process the word and object and more rapidly shift attention to the right object (Fernald, Zangl, Portillo, & Marchman, 2008). Other studies manipulating the sound properties of the words have shown in the time course of looking at two potential referents, that the time course of competition and its resolution depend on the similarity of cues (Swingley & Aslin, 2007). Other recent studies showed that time course of looking to the named referent closely tracked the co-occurrence probabilities of words, referents, and

distracters (Fernald, Thorpe, & Marchman, 2010; Vouloumanos & Werker, 2009; Yu & Smith, in press). These finer-grained methods might also be used to directly test the idea that learned clusters of coherently varying cues organize orienting to, detecting, and processing of predicted properties in ways that nimbly guide attention to the right components of a scene for word learning.

We also need studies on cued attention outside of the domain of early word learning, studies that investigate the processes and mechanisms in adult literature. At present there are very few studies (Kirkham, Slemmer, & Johnson, 2002; Reid, Striano, Kaufman, & Johnson, 2004; Richards, 2005; Sloutsky & Robinson, 2008; Wu & Kirkham, in press) and no active area of concentrated research effort. The studies that do exist suggest that cued attention will not work exactly the same way in children as in adults. At present, we know very little about attentional learning and cued attention in young children. One distinction in the study of attentional processes that may be particularly relevant is that between exogenous and endogenous cueing (see Colombo, 2001; Goldberg et al., 2001; Smith & Chatterjee, 2008; Snyder & Munakata, 2008, 2010). Exogenous attention refers to the quick capture of attention by salient stimuli, such as the flashing of lights. Endogenous attention is attention directed by associated cues. Many studies of endogenous versus exogenous cueing use the covert-orienting paradigm (Jonides, 1980; Posner & Raichle, 1994): A cue directs attention to the target information either exogenously (a flashing light) or endogenously (by a previous association) and either correctly or incorrectly. There is considerable development in the dynamics of both exogenous and endogenous cueing and in disengaging attention when miscued (see Colombo, 2001; Goldberg et al., 2001). However, advances in endogenous cueing appear to lag behind exogenous cueing particularly in the overriding of misleading cues and in switching attention given a new cue (Cepeda, Kramer, & de Sather, 2001; Snyder & Munakata, 2010). By one proposal the major developmental changes lie in the processes that resolve cue competition (Snyder & Munakata, 2008, 2010). Two recent studies of preschoolers in attention-switching tasks suggest that experimentally training clusters of different cues for targets (Sloutsky & Fisher, 2008) or training more abstract cues (Snyder & Munakata, 2010) enhances switching, a result that may link cued attention, word learning, and executive control of attention.

In brief, internally directed attention that smartly moves to the right information for learning is likely to have its own compelling developmental story, with successful attention dependent on one's history with a particular set of cues and overlapping mechanisms of activation, preactivation, cue interactions, and competition. We need systematic and fine-grained studies of the development of cued attention from infancy through the early word-learning period and into childhood.

8. Knowledge as process

The systematicity with which young children learn the concepts they share with their elders, and they systematicity with which they generalize a newly learned name in different ways for different kinds, clearly demonstrates that they have knowledge. That is not in

doubt; the contentious question is the nature of the knowledge. Macro-level accounts of this knowledge have been in terms of propositions and core principles about kinds of categories and kinds of word meanings, knowledge of the sort proposed in the 1990 special issue (see also Booth & Waxman, 2002; Dewar & Xu, 2009; Diesendruck & Bloom, 2003; Hall, 1996; Soja, 1992). These explanations are based on the traditional view of cognition as sharply divided into knowledge and process and competence and performance. Thus, in terms of this traditional framework, the two levels of knowledge and process illustrated on the left of Fig. 7 were separable and one could sensibly ask of children's performance in any word-learning task whether their performance reflected knowledge about how words map to meaning (competence) or whether it reflected some performance variable. One example of this are the various debates as to whether the shape bias emerges because of conceptual knowledge about how nouns map to categories (e.g., Booth & Waxman, 2002; Cimpian & Markman, 2005) or whether it emerges because of attentional learning (Smith & Samuelson, 2006). This traditional explanation would seem to suggest that one could pit conceptual explanations against explanations in terms of (only) real-time processing. This framing of the question—concepts versus attention—is problematic on its own terms. In any real-time

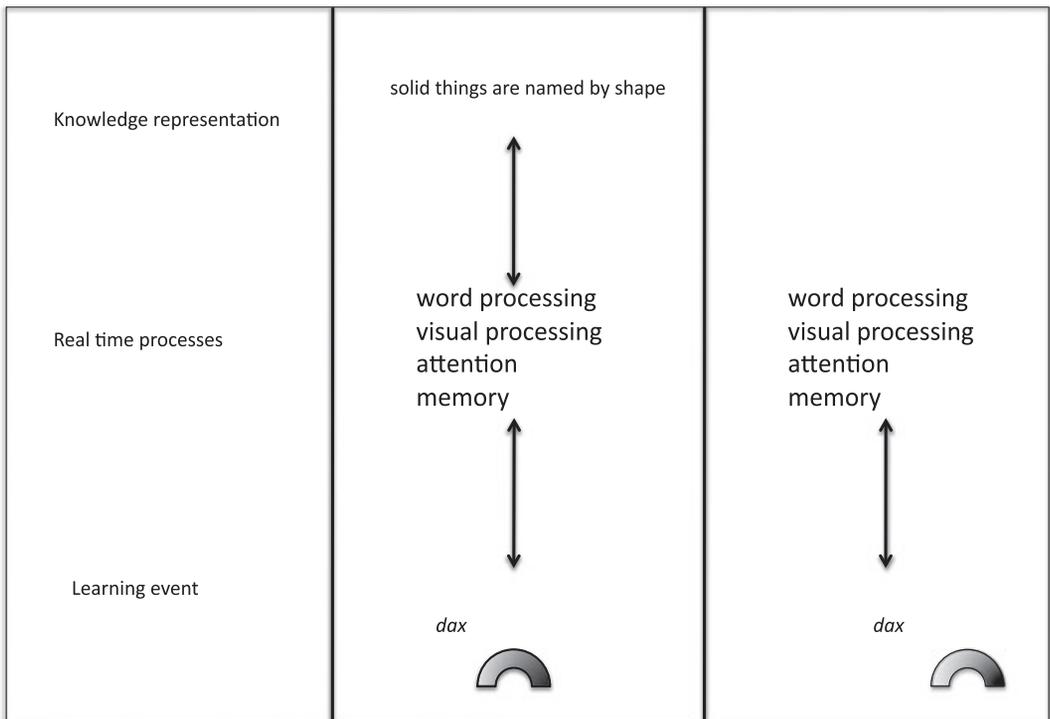


Fig. 7. An illustration of two views of cognition: On the left, knowledge representations are connected to real-time task performance as separable aspects of cognition. On the right, processes (which through their operating characteristics and history instantiate knowledge) generate performance in real-time tasks. Within this framework, beliefs, concepts, and propositional knowledge are higher level theoretical descriptions of the knowledge embedded in these processes.

task, conceptual knowledge requires supporting processes as well—the perceptual, attention, memory, and decision processes. Whatever can be explained by the account on the right (the process-only account) can *necessarily* be explained by the account on the left: Whatever can be explained by A (process) can *always* also be explained by A+B (process plus propositions).

The framing of the contending explanations in terms of concepts versus processes such as attention is also more profoundly misguided. Modern-day understanding of neural processes makes clear that knowledge has no existence outside of process and that the relevant processes often encompass many different systems and time-scales (see Barsalou, 2009). What we call knowledge is an abstraction over many underlying processes. Although these higher level abstractions may summarize meaningful regularities in the operating characteristics of the system as a whole, we also need to go underneath them—and understand the finer-grained processes from which they are made. “Word learning rules” and “concepts” may be to attentional processes as “hunger” is to hormones: Not a separate mechanism but a portmanteau abstraction that includes attentional learning (as well as other processes). Clearly, one goal of science is to unpack these carry-all abstractions. Cued attention is a particularly intriguing potential mechanism from this perspective because it is one that melds competence and performance, knowledge and process, perception and conception. It does so because cued attention is a process, operating in real time, strongly influenced by the momentary input, with effects at the sensory and perceptual level but also driven by the rich data structure of predictive relations between cues and outcomes in a lifetime of experiences. As such, attention is a process that aggregates, acquires, and applies knowledge.

Acknowledgments

The research summarized in this paper was supported by NIMH grant R01MH60200, and NICHD grants R01HD 28675 to Linda Smith and NICHD 1R01HD058620-01 to Hanako Yoshida.

References

- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of Royal Society B*, 364, 1281–1289.
- Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research*, 49, 1154–1165.
- Bichot, N. P., & Rossi, A. F. (2005). Parallel and serial neural mechanisms for visual search in Macaque Area V4. *Science*, 308(5721), 529–534.
- Bichot, N. P., Schall, J., & Thompson, K. (1996). Visual feature selectivity in frontal eye fields induced by experience in mature macaques. *Nature*, 381, 697–699.
- Biederman, I. (1972). Perceiving real-world scenes. *Science*, 177(4043), 77–80.
- Billman, D., & Heit, E. (1989). Observational learning from internal feedback: A simulation of an adaptive learning method. *Cognitive Science*, 12, 587–625.

- Billman, D., & Knutson, J. (1996). Unsupervised concept learning and value systematicity: A complex whole aids learning the parts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(2), 458–475.
- Bлага, O. M., & Colombo, J. (2006). Visual processing and infant ocular latencies in the overlap paradigm. *Developmental Psychology*, 42(6), 1069–1076.
- Booth, A. E., & Waxman, S. R. (2002). Word learning is ‘smart’: Evidence that conceptual information affects preschoolers’ extension of novel words. *Cognition*, 84, B11–B22.
- Boroditsky, L. (2001). Does language shape thought? English and Mandarin speakers’ conceptions of time. *Cognitive Psychology*, 43, 1–22.
- Bott, L., Hoffman, A., & Murphy, G. (2007). Blocking in category learning. *Journal of Experimental Psychology: General*, 136, 685–699.
- Bowerman, M. (1996). Cognitive versus linguistic determinants. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 145–186). Cambridge, England: Cambridge University Press.
- Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Interfering neighbours: The impact of novel word learning on the identification of visually similar words. *Cognition*, 97, B45–B54.
- Boyce, S. J., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 556–566.
- Boynton, G. M. (2009). A framework for describing the effects of attention on visual responses. *Vision Research*, 49, 1129–1143.
- Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology*, 33, 798–815.
- Cepeda, N. J., Kramer, A. F., & de Sather, J. (2001). Changes in executive control across the life span: Examination of task switching performance. *Developmental Psychology*, 37, 715–730.
- Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition*, 18(5), 537–545.
- Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 106(14), 5529–5533.
- Chelazzi, L. et al. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, 363, 345–347.
- Chelazzi, L. et al. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, 80, 2918–2940.
- Cheng, P. W., & Holyoak, K. J. (1995). Complex adaptive systems as intuitive statisticians: Causality, contingency, and prediction. In H. L. Roitblat & J. Meyer (Eds.), *Comparative approaches to Cognitive Science* (pp. 271–302). Cambridge, MA: The MIT Press.
- Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, 41, 83–121.
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36(1), 28–71.
- Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current opinion in Neurobiology*, 17, 177–184.
- Cimpian, A., & Markman, E. (2005). The absence of a shape bias in children’s word learning. *Developmental Psychology*, 41, 1003–1019.
- Clohessy, A. B., Posner, M. I., & Rothbart, M. K. (2001). Development of the functional visual field. *Acta Psychologica*, 106, 51–68.
- Colombo, J. (2001). The development of visual attention in infancy. In S. T. Fiske, D. L. Schacter, & C. Zahn-Waxler (Eds.), *Annual review of psychology* (pp. 337–367). Palo Alto, CA: Annual Reviews.
- Colunga, E. (2006). The effect of priming on preschooler’s extensions of novel words: How far can “dumb” processes go? *Proceedings of the 30th Annual Boston University Conference on Language Development*, 23, 96–106.

- Colunga, E., & Smith, L. B. (2004). Dumb mechanisms make smart concepts. *Proceedings of the Annual Conference of the Cognitive Science Society*, 26, 239–244.
- Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*, 112(2), 347–382.
- Colunga, E., & Smith, L. B. (2008). Flexibility and variability: Essential to human cognition and the study of human cognition. *New Ideas in Psychology*, 26, 174–192.
- Colunga, E., Smith, L. B., & Gasser, M. (2009). Correlation versus prediction in children's word learning: Cross-linguistic evidence and simulations. *Language and Cognition*, 1(2), 197–217.
- Davis, C. J., & Lupker, S. J. (2006). Masked inhibitory priming in English: Evidence for lexical inhibition. *Journal of Experimental Psychology: Human Perception and Performance*, 32(3), 668–687.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Reviews of Neuroscience*, 18, 193–222.
- Dewar, K., & Xu, F. (2009). Do early nouns refer to kinds or distinct shapes? Evidence from 10-month-old infants. *Psychological Science*, 20, 252–257.
- Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70–95). New York: Oxford University Press.
- Diesendruck, G., & Bloom, P. (2003). How specific is the shape bias? *Child Development*, 74, 168–178.
- Duncan, J. (1996). Cooperating brain systems in selective perception and action. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI* (pp. 549–578). Cambridge, MA: The MIT Press.
- Ellis, N. C. (2006). Selective attention and transfer phenomena in LZ acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27 (2), 164–194.
- Fernald, A., Thorpe, K., & Marchman, V. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective–noun phrases. *Cognitive Psychology*, 60, 190–217.
- Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. Sekerina, E. Fernández, & H. Clahsen (Eds.), *Language processing in children* (pp. 97–135). Amsterdam, The Netherlands: Benjamins.
- Gathercole, V., Cramer, L., Somerville, S., & Jansen op de Haar, M. (1995). Ontological categories and function: Acquisition of new names. *Cognitive Development*, 10, 225–251.
- Gathercole, V., & Min, H. (1997). Word meaning biases or language-specific effects? Evidence from English, Spanish and Korean. *First Language*, 17(49 Pt 1), 31–56.
- Gathercole, V., Thomas, E. M., & Evans, D. (2000). What's in a noun? Welsh-, English-, and Spanish-speaking children see it differently. *First Language*, 20, 55–90.
- Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and animate-inanimate distinction as examples. *Cognitive Science*, 14, 79–106.
- Gelman, S. A., & Coley, J. D. (1991). Language and categorization: The acquisition of natural kind terms. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 146–196). New York: Cambridge University Press.
- Gentner, D., & Boroditsky, L. (2001). Individuation, relational relativity and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge, England: Cambridge University Press.
- Gilbert, C., Ito, M., Kapadia, M., & Westheimer, G. (2000). Interactions between attention, context and learning in primary visual cortex. *Vision Research*, 40, 1217–1226.
- Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review Neuroscience*, 30, 535–574.
- Goldberg, M., Maurer, D., & Lewis, T. (2001). Developmental changes in attention: The effects of endogenous cueing and of distractors. *Developmental Science*, 4, 209–219.
- Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, 49, 585–612.

- Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, 21(1), 125–155.
- Grossberg, S. (1982). Processing of expected and unexpected events during conditioning and attention: A psychophysiological theory. *Psychological Review*, 89(5), 529–572.
- Grossmann, T., & Farroni, T. (2009). Decoding social signals in the infant brain: A look at eye gaze perception. In M. de Haan & M. R. Gunnar (Eds.), *Handbook of developmental social neuroscience* (pp. 87–106). New York: Guilford Press.
- Halberda, J. (2006). Is this a dax which I see before me? use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, 53(4), 310–344.
- Hall, D. G. (1996). Naming solids and non-solids: Children's default construals. *Cognitive Development*, 11, 229–264.
- Hanania, R., & Smith, L. B. (in press). Selective attention and attention switching: Towards a unified developmental approach. *Developmental Science*.
- Hirotsani, M., Stets, M., Striano, T., & Friederici, A. D. (2009). Joint attention helps infants learn new words: Event-related potential evidence. *NeuroReport: For Rapid Communication of Neuroscience Research*, 20(6), 600–605.
- Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R., Brand, R. J., Brown, E., Chung, H. L., Hennon, E., & Rocroi, C. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the society for Research in Child Development*, 65(3), v–123.
- Horst, J., Scott, E., & Pollard, J. (in press). The role of competition in word learning via referent selection. *Developmental Science*.
- Iannucci, J. E. (1952). *Lexical number in Spanish nouns with reference to their English equivalents*, Philadelphia: University of Pennsylvania.
- Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition*, 62, 169–200.
- James, W. (1980). *The Principles of psychology*. Chicago: University of Chicago Press.
- Jiang, Y., & Chun, M. M. (2001). Asymmetric object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, 27(4), 895–918.
- Jiang, Y., & Chun, M. M. (2003). Contextual cueing: Reciprocal influences between attention and implicit learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 277–296). Amsterdam, The Netherlands: John Benjamins Publishing Company.
- Jiang, Y., & Leung, A. W. (2005). Implicit learning of ignored visual context. *Psychonomic Bulletin & Review*, 12(1), 100–106.
- Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 683–702.
- Jones, S. S., & Smith, L. B. (1998). How children name objects with shoes. *Cognitive Development*, 13, 323–334.
- Jones, S. S., & Smith, L. B. (2002). How children know the relevant properties for generalizing object names. *Developmental Science*, 5, 219–232.
- Jones, S. S., Smith, L. B., & Landau, K. B. (1991). Object properties and knowledge in early lexical learning. *Child Development*, 62, 499–516.
- Jonides, J. (1980). Towards a model of the mind's eye's movement. *Canadian Journal of Psychology*, 34, 103–112.
- Kamin, L. J. (1968). 'Attention-like' processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9–33). Coral Gables, FL: University of Miami Press.
- Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment* (pp. 279–296). New York: Appleton-Century-Crofts.
- Kastner, S. et al. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, 22, 751–761.

- Keil, F. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 234–254). New York: Cambridge University Press.
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83, B35–B42.
- Kobayashi, H. (1998). How 2-year-old children learn novel part names of unfamiliar objects. *Cognition*, 68, B41–B51.
- Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, 99, 22–44.
- Kruschke, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 3–26.
- Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology*, 45(6), 812–863.
- Kruschke, J. K. (2005). Learning involves attention. In G. Houghton (Ed.), *Connectionist models in cognitive psychology*, Ch. 4 (pp. 113–140). Hove, East Sussex, UK: Psychology Press.
- Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin & Review*, 7(4), 636–645.
- Kruschke, J. K., & Johansen, M. K. (1999). A model of probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(5), 1083–1119.
- Kruschke, J. K., Kappenman, E. S., & Hetrick, W. P. (2005). Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 31, 830–845.
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, 3, 299–321.
- Landau, K. B., Smith, L. B., & Jones, S. (1998). Object perception and object naming in early development. *Trends in Cognitive Science*, 2, 19–24.
- Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge, England: Cambridge University Press.
- Lewicki, P., Hill, T., & Czyzewska, M. (1992). Nonconscious acquisition of knowledge. *American Psychologist*, 47(6), 796–801.
- Lewicki, P., Hill, T., & Czyzewska, M. (1997). Hidden covariation detection: A fundamental and ubiquitous phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(1), 221–228.
- Lubow, R. E. (1997). Latent inhibition as a measure of learned inattention: Some problems and solutions. *Behavioural Brain Research Special Issue: Psychobiology of Learned Inattention*, 88(1), 75–83.
- Lubow, R. E., & Kaplan, O. (1997). Visual search as a function of type of prior experience with target and distractor. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1), 14–14.
- Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: The effect of nonreinforced preexposure to the conditioned stimulus. *Journal of Comparative and Physiological Psychology*, 52, 415–419.
- Ludwig, C. J. H., Gilchrist, I. D., & McSorley, E. (2005). The remote distractor effect in saccade programming: Channel interactions and lateral inhibition. *Vision Research*, 45(9), 1177–1190.
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82, 276–298.
- Mackintosh, N. J. (1976). Overshadowing and stimulus intensity. *Animal Learning & Behavior*, 4, 186–192.
- Mackintosh, N. J., & Turner, C. (1971). Blocking as a function of novelty of CS and predictability of UCS. *Quarterly Journal of Experimental Psychology*, 23, 359–366.
- Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Series in learning, development, and conceptual change. Cambridge, MA: MIT Press.
- Markman, E. M. (1990). Constraints children place on word meaning. *Cognitive Science*, 14, 57–77.
- Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neuroscience*, 29(6), 317–322.

- Mayor, J., & Plunkett, K. (2010). A neurocomputational account of taxonomic responding and fast mapping in early word learning. *Psychological Review*, *117*(1), 1–31.
- McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Review Neuroscience*, *4*(4), 310–322.
- McPherson, L. (1991). A little goes a long way: Evidence for a perceptual basis of learning for the noun categories COUNT and MASS. *Journal of Child Language*, *18*, 315–338.
- Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *8*(1), 37–50.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207–238.
- Miller, J. L., Ables, E. M., King, A. P., & West, M. J. (2009). Different patterns of contingent stimulation differentially affect attention span in prelinguistic infants. *Infant Behavior & Development*, *32*(3), 254–261.
- Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. *Cognition*, *84*(3), 267–293.
- Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, *9*(3), 273–302.
- O'Reilly, R. C. (2001). Generalization in interactive networks: The benefits of inhibitory competition and Hebbian learning. *Neural Computation*, *13*, 1199–1242.
- Pessoa, L., Kastner, S., & Ungerleider, L. G. (2003). Neuroimaging studies of attention: From modulation of sensory processing to top-down control. *The Journal of Neuroscience*, *23*(10), 3990–3998.
- Posner, M., & Raichle, M. (1994). *Images of mind*. New York: W. H. Freeman and Company.
- Posner, M., & Rothbart, M. K. (2007). Research on attention as a model for the integration of psychological science. *Annual Review of Psychology*, *58*, 1–23.
- Posner, M. I., Rothbart, M. K., Thomas-Thrapp, L., & Gerardi, G. (1998). The development of orienting to locations and objects. In R. D. Wright (Ed.), *Visual attention* (pp. 269–288). New York: Oxford University Press.
- Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (in press). Feature-label-order effects and their implications for symbolic learning. *Cognitive Science*.
- Regier, T. (2005). The emergence of words: Attentional learning in form and meaning. *Cognitive Science: A Multidisciplinary Journal*, *29*(6), 819–865.
- Reid, V. M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cueing facilitates neural processing of objects in 4-month-old infants. *NeuroReport: For Rapid Communication of Neuroscience Research*, *15*(16), 2553–2555.
- Rescorla, R. A., & Wagner, A. R. (1972) *A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement*. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II* (pp. 64–99). New York: Appleton-Century-Crofts.
- Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, *61*, 168–185.
- Richards, J. E. (2005). Localizing cortical sources of event-related potentials in infants' covert orienting. *Developmental Science*, *8*(3), 255–278.
- Richards, J. E. (2008). Attention in young infants: A developmental psychophysiological perspective. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (2nd ed.) pp. 479–497). Cambridge, MA: MIT Press.
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.
- Rossi, A. F., & Paradiso, M. A. (1995). Feature-specific effects of selective visual attention. *Vision Research*, *35*(5), 621–634.
- Samuelson, L. (2002). Statistical regularities in vocabulary guide language acquisition in connectionist models and 15-20-month-olds. *Developmental Psychology*, *38*, 1016–1037.
- Samuelson, L. K., Horst, J. S., Schutte, A. R., & Dobbertin, B. (2008). Rigid thinking about deformables: Do children sometimes overgeneralize the shape bias? *Journal of Child Language*, *35*, 559–589.

- Samuelson, L. K., Schutte, A. R., & Horst, J. S. (2009). The dynamic nature of knowledge: Insights from a dynamic field model of children's novel noun generalizations. *Cognition*, *110*, 322–345.
- Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure, and syntax correspond? *Cognition*, *73*(1), 1–33.
- Samuelson, L. K., & Smith, L. B. (2000a). Children's attention to rigid and deformable shape in naming and no naming tasks. *Child Development*, *71*(6), 1555–1570.
- Samuelson, L. K., & Smith, L. B. (2000b). Grounding development in cognitive processes. *Child Development*, *71*, 98–106.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66.
- Shanks, D. R. (1985). Forward and backward blocking in human contingency judgement. *Quarterly Journal of Experimental Psychology*, *37B*, 1–21.
- Shanks, D. R., Channon, S., Wilkinson, L., & Curran, H. V. (2006). Disruption of sequential priming in organic and pharmacological amnesia: A role for the medial temporal lobes in implicit contextual learning. *Neuropharmacology*, *31*(8), 1768–1776.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*(2), 127–190.
- Sloutsky, V. M., & Fisher, A. V. (2008). Attentional learning and flexible induction: How mundane mechanisms give rise to smart behaviors. *Child Development*, *79*, 639–651.
- Sloutsky, V. M., & Robinson, C. W. (2008). The role of words and sounds in infants' visual processing: From overshadowing to attentional tuning. *Cognitive Science: A Multidisciplinary Journal*, *32*(2), 342–365.
- Smith, L. B. (1995). Self-organizing processes in learning to learn words: Development is not induction. In C. Nelson (Ed.), *Basic and applied perspectives on learning, cognition, and development*, Vol. 28 (pp. 1–32). Mahwah, NJ: Lawrence Erlbaum Associates.
- Smith, S. E., & Chatterjee, A. (2008). Visuospatial attention in children. *Neurological Review*, *65*, 1284–1288.
- Smith, L. B., Colunga, E., & Yoshida, H. (2003). Making an ontology: Cross-linguistic evidence. In D. Rakison & L. Oakes (Eds.), *Early category and concept development: Making sense of the blooming, buzzing confusion* (pp. 275–302). London: Oxford University Press.
- Smith, L. B., & Gasser, M. (2005). The development of embodied cognition: Six lessons from babies. *Artificial Life*, *11*, 13–39.
- Smith, L. B., Jones, S. S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretation. *Developmental Psychology*, *28*, 273–286.
- Smith, L. B., Jones, S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*(1), 13–19.
- Smith, L. B., & Samuelson, L. (2006). An attentional learning account of the shape bias: Reply to Cimpian and Markman (2005) and Booth, Waxman, and Huang (2005). *Developmental Psychology*, *42*, 1339–1343.
- Snyder, H. R., & Munakata, Y. (2008). So many options, so little time: The roles of association and competition in underdetermined responding. *Psychological Bulletin & Review*, *15*, 1083–1088.
- Snyder, H. R., & Munakata, Y. (2010). Becoming self-directed: Abstract representations support endogenous flexibility in children. *Cognition*.
- Soja, N. N. (1992). Inferences about the meaning of nouns; the relationship between perception and syntax. *Cognitive Development*, *7*, 29–45.
- Soja, N. N. (1994). Evidence for a distinct kind of noun. *Cognition*, *51*(3), 267–284.
- Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meanings: Object terms and substance terms. *Cognition*, *38*(2), 179–211.
- Spitzer, H. et al. (1988). Increased attention enhances both behavioral and neuronal performance. *Science*, *240*, 338–340.
- Summerfield, C., Egnor, T., Greene, M., Koehlin, E., Mangels, J., & Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science*, *314*(5803), 1311–1314.

- Swingle, D., & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, *54*, 99–132.
- Treue, S., & Martinez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, *399*, 575–579.
- Vouloumanos, A., & Werker, J. F. (2009). Infants' learning of novel words in a stochastic environment. *Developmental Psychology*, *45*, 1611–1617.
- Walley, R. E., & Weiden, T. D. (1973). Lateral inhibition and cognitive masking: A neuropsychological theory of attention. *Psychological Review*, *80*(4), 284–302.
- Ward, T. B., Becker, A. H., Hass, S. D., & Vela, E. (1991). Attribute availability and the shape bias in children's category generalization. *Cognitive Development*, *6*(2), 143–167.
- Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories. *Cognitive Psychology*, *29*(3), 257–302.
- Williford, T., & Maunsell, J. H. (2006). Effects of spatial attention on contrast response functions in macaque area V4. *Journal of Neurophysiology*, *96*, 40–54.
- Wu, R., & Kirkham, N. (in press). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology*.
- Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, *396*, 72–75.
- Yoshida, H., & Hanania, R. (2007). Attentional highlighting as a mechanism behind early word learning. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th annual meeting of the Cognitive Science Society* (pp. 719–724). Austin, TX: Cognitive Science Society.
- Yoshida, H., & Smith, L. B. (2001). Early noun lexicons in English and Japanese. *Cognition*, *82*, 63–74.
- Yoshida, H., & Smith, L. B. (2003a). Shifting ontological boundaries: How Japanese- and English- speaking children generalize names for animals and artifact. *Developmental Science*, *6*(1), 1–34.
- Yoshida, H., & Smith, L. B. (2003b). Correlation, concepts and cross-linguistic differences. *Developmental Science*, *6*(1), 30–34.
- Yoshida, H., & Smith, L. B. (2005). Linguistic cues enhance the learning of perceptual cues. *Psychological Science*, *16*(2), 90–95.
- Yoshida, H., Swanson, J., Drake, C., & Gudel, L. (2001). Japanese- and English-speaking children's use of linguistic cues to animacy in category formation. Biennial meeting of the Society for Research on Child Development, Minneapolis, MN.
- Younger, B. A., & Cohen, L. B. (1983). Infant perception of correlations among attributes. *Child Development*, *54*, 858–867.
- Yu, C., & Smith, L. B. (in press). What you learn is what you see: Using eye movements to study infant cross-situational word learning. *Developmental Science*.
- de Zubicaray, G. I., & McMahon, K. L. (2009). Auditory context effects in picture naming investigated with event-related fMRI. *Cognitive, Affective & Behavioral Neuroscience*, *9*(3), 260–269.