
LINDA B. SMITH: Avoiding Associations When It's Behaviorism You Really Hate

"Association of ideas" is an unfashionable term—much used in the 19th century, it has almost disappeared in this one—but the time may have come to put it back to work again. The reason for its disappearance was the doubt, earlier, whether ideas and thought exist at all. We have seen that Thorndike raised this question about cats; John B. Watson took it further, and suggested that human thinking is only a series of tiny muscular contractions, each contraction providing a stimulus to the next one, so that in thinking one is really talking to oneself under one's breath. In that case ideas would not exist. . . . We now know however that ideas do exist . . . and we also know that connections can be set up between them. (D. O. Hebb, *Textbook of Psychology*)

The chapters in this volume present wonderful insights into early word learning. Many of the authors remark on babies' attention to and learning about relations—relations between forms and referents, between speakers' actions and speakers' intentions, between direction of gaze and objects, between syntactic cues and meanings. They use such words as "links," "maps," "correspondence," "predicts," and "expectations" in their discussions of what babies know about learning words. They avoid references to "associates" and "association." Among cognitive developmentalists "association" is still a very unfashionable term—a term to be raised only if the purpose is to dismiss it.

My fellow contributors to this volume offer three reasons for running from associative accounts: Associative learning implies that (1) contiguity is everything, (2) we can learn anything, and (3) we have, more or less, the mental lives of rats. They are wrong. Associative learning implies none of this. My fellow authors confuse associative learning with the worst of American behaviorism. As Hebb (1949) notes in his history, associative learning was on the other side of the behaviorist firestorm. It went underground in the behaviorist heyday and was at the vanguard of the cognitive revolution (e.g., Hebb; Minsky & Papert, 1969). Radical behaviorists hated the idea of associations. Why? Because associative learning is about ideas, about internal mental events, about the processes that *make* cognition.

Here, then, is the take-home message: Today's associative learning is not your father's behaviorism. Associative mechanisms do not imply what developmentalists take them to imply. Indeed, way too much is known about associative processes—from neuroscience, from behavioral experiments, from mathematical and computational modeling—for reflexive rejections to have a place in developmental discussions of mechanism any longer (see also Kelly & Martin, 1994). Although there is not enough space in this brief commentary to provide even a partial tutorial, I present three brief lessons about associations, each mo-

tivated by my colleagues' confusions. I encourage readers to pursue a more thorough study of the field.

Three Lessons

Predictability and Causality. Classical conditioning is a basic form of associative learning. It consists of learning the relation between pairs of sensory events, one of which, the unconditioned stimulus (UCS), has a preexisting connection to a measurable response. Classical conditioning is not the only form of associative learning, and it is *not* a candidate model for word learning. Nonetheless, it is useful to consider just what knowledge organisms acquire in this most rudimentary form of learning. What is learned is the predictive relation between the conditioned stimulus, or CS (the neutral stimulus), and the UCS (the consequential sensory event), *whatever that predictive relation is*. If the CS reliably precedes the UCS by a short time, organisms learn that the CS predicts that the UCS will occur after a short delay; if the CS reliably precedes the UCS by a long time, organisms learn that the CS predicts that the UCS will occur after a long delay. If the UCS precedes the CS (backward conditioning), organisms learn not to expect the UCS to occur immediately after the CS occurs. In brief, the evidence of the past three-quarters of a century shows that just about any set pattern of paired experiences leads to learning and what, precisely, is learned is the predictive relation between the cues. Indeed, classical conditioning is so much about making predictions, about expectations, that the Rescorla and Wagner model (1972) of classical conditioning forms the basis for several successful formal models of human causal reasoning (see Cheng & Holyoak, 1995). Here is the lesson: *Associative learning, even in its most rudimentary form, is neither about contiguity nor about counting the pairings of events in experience; it is about prediction and forming expectations.*

Attention and Constraints. Because attention is very much at the heart of all forms of learning, it is a centerpiece of several of the chapters in this volume. Simply, we learn about what we attend to, and we learn *what to attend to*. Over the past quarter century, considerable theoretical and empirical work in cognitive psychology has yielded a nearly unified theory of the associative basis of attentional learning. The foundational ideas are those of Rescorla and Wagner (1972); Mackintosh (1975); Shepard, Hovland, and Jenkins (1961); and Medin and Schaffer (1978). The mathematical unification of these ideas into a single theory is gaining momentum through the work of Nosofsky (1986) and Kruschke (1992, 1999). The central idea is simple but powerful: As organisms learn associations among cues and outcomes, they also learn which cues to attend to. The power of these ideas as realized in current formal theories is evident in the success of these theories in explaining (in fine experimental de-

tail) phenomena from many divergent fields and paradigms. The associative mechanisms behind attentional learning instantiated in, for example, Kruschke's models of adult category learning (e.g., 1992, 1999) constitute the accepted explanation of attentional learning in psychology—at least they do outside of developmental psychology. There is simply no equally specified or empirically supported contender of how organisms—including people—shift attention to task-relevant properties. Thus, for *any* of the attentional phenomena that are part of early word learning, attentional learning via associative mechanisms is, *prima facie*, a candidate explanation, one to be dismissed only when an *equally specified* alternative is offered that can explain and unify as many phenomena.

Science seeks unified explanations. It is important, then, that attentional learning via associative mechanisms may explain word-learning phenomena beyond the shape bias. Consider the recent finding that words become special in their ability to refer (Namy & Waxman, 1998; Woodward & Hoyne, 1999). Recall that Namy and Waxman presented 18- and 26-month-olds with a triad of objects: an exemplar and two choice objects. In one condition, the experimenter named the exemplar object with a novel name and, using that name, asked the child to select among the two choice objects. In the second condition, the experimenter referred to the exemplar with a hand gesture rather than with a spoken name and, using the gesture, asked the child to select among the two choice objects. The younger children, 18-month-olds, chose taxonomically in both conditions. Apparently, for younger children, any associate of an object can work to push attention to similar kinds. The older children, in contrast, chose taxonomically only in the name condition; they responded randomly when signaled to make a choice by a gesture.

Some contributors to this volume have suggested that this fact shows that, at least for older children, word learning cannot be *merely* associative learning. If it were, they reason, the older children would associate the gesture with the object. This reasoning is flawed, however, because associative learning does not imply that any cue can be linked with any other cue. Indeed, a powerful (and well-studied) phenomenon in associative learning—blocking—provides a mechanistic explanation of Namy and Waxman's (1998; and Woodward & Hoyne's, 1999) results. Blocking is the name given to the following powerful fact about associative learning: If one already has learned a cue that reliably predicts an outcome, it is harder to learn a *new* cue that predicts the same outcome (see Mackintosh, 1975; Kruschke, 1999). In terms of Namy and Waxman's result, an already learned link between naming and the object of the speaker's attention *blocks* the learning of a link between a gesture and the object of the speaker's attention. Think of what the mere fact of blocking means: Associative learning based on the regularities in one's previous experiences *alters* what regularities will be learned in the future.

Associative learning is not the creation of isolated connections between

stimulus events. The formation of initially simple associations *changes* what is attended to and, in so doing, changes what will be learned in the future. This is where the significance of a learned shape bias lies. And here is the lesson: *Learning is a historical process that leads to increasingly constrained destinies. Unbiased associative mechanisms become biased learning mechanisms.*

Universality and Specificity. A number of phenomena, such as blocking, have been observed in a variety of learning domains and in a number of species (see Kruschke, 1999). This commonality suggests similar learning principles and a similarity of mechanisms both across domains within a single species and across different species. This does *not* constitute a weakness—a fatal flaw—in associative learning as a candidate mechanism for children’s word learning. The argument that associative learning cannot be the basis of children’s word learning because rats do not learn words and the related argument that associative learning, *if* involved, is unimportant because it cannot explain the differences between people and rats are both profoundly wrongheaded. Using associative mechanisms to explain word learning in children no more implies that rats should be able to learn words than the universality of the DNA code implies that we should have rat bodies. And the fact that the DNA code for protein synthesis is universal despite the fact that rats and people have very different bodies does not make the DNA code ipso facto uninteresting and unimportant in explaining either the differentiation of organs within a single body or bodily differences across species. In both cases—the same DNA code or the same associative mechanisms—the uniqueness of the outcome depends exquisitely on the particulars in the *history* of the processes and the cascading consequences within that history. In other words, what matters is development. I offer two striking examples of associative learning in developmental process.

The first example concerns suckling behavior in newborn rats. Shortly after birth and despite the severe immaturity of the sensory and motor system, the newborn rat moves itself to its mother’s ventrum and suckles. This behavior—crucial to survival—is created by a complex chain of events that includes classical conditioning. First, research has shown that the scent of amniotic fluid on the mother’s ventrum (deposited when the mother licks the pups and then herself after birth) is a crucial cue. Newborn rats fail to attach to nipples if the ventrum and nipples are washed (Teicher & Blass, 1976, 1977). Second, this scent cue is learned. Pedersen and Blass (1982) introduced an arbitrary smell (lemon) into the amniotic sac two days before birth. In this condition, rat pups did not suckle in the context of amniotic fluid but did so in the presence of a lemon scent. Third, recent research shows that the scent cue gains its power to initiate suckling through classical conditioning (Ronca, Abel, & Alberts, 1996; Abel, Ronca, & Alberts, 1998). The CS, the smell, is linked to the UCS, compression, through the birthing process. Abel, Ronca, and Alberts term this form of peri-

natal conditioning “intrinsic,” because the associations learned are derived from the functional properties of birth. It is by virtue of the natural sequence of concordant events constituting the birth process that the conditioning occurs and that the congenital capability of suckling so crucial to survival appears.

The second example of how universal mechanisms lead to specific (and constrained) outcomes concerns imprinting and, in particular, a model of imprinting offered by O’Reilly and Johnson (1994). O’Reilly and Johnson used a recurrent network composed of the most general and universal processes: Hebbian learning, excitatory connections, and lateral inhibition. None of these is a prescription for imprinting or for its sensitive period, and each is evident everywhere in the study of brains of all species—the “common stuff” of brains. The experiences that O’Reilly and Johnson gave the network are also “common stuff”—looking at individual objects presented one at a time for varying durations, experiences that might correspond to 10 minutes or 1 hour of viewing the same object. That is all there is to the model. But the outcome of these experiences is preferential recognition of the first object seen and the emergence of a self-terminating sensitive period. Quite simply, if early experience consists of an object that persists for sufficient duration, the strength of that bias cannot ever be overcome; through lateral inhibition and recurrent connections, it maintains itself. If, in contrast, early experience consists of sufficiently many different nonpersisting objects, no preference emerges and the possibility of developing such a preference is lost. The point of this example—as in the example of suckling in baby rats and as in the associative-learning account of the shape bias—is that something special can develop out of the particular history of activity of general learning processes that subserve other specializations.

Here, then, is the third lesson: *Associative learning operates in particular contexts, as parts of particular causal chains and particular natural histories, and yields specific outcomes.*

A Note about Levels of Explanation

These three lessons are profoundly important for evaluating the different accounts in this volume. They mean that it is possible to build an associative-learning version of each of the alternative explanations—the social-pragmatic, the constraints, the many-principles, the active-child accounts. These associative-learning accounts would be implementation versions of their parent explanations, versions that mechanistically specify the undefined terms of “links” and “maps” and “predicts” and “expects” that fill the parent accounts. Specifying such implementation versions of the other accounts in this volume is a useful goal, since each account captures real and important truths about children; about language; and about how children become, in such a very short time, truly prodigious learners of words. But, at present, each is currently underspecified,

Similarly, Clark's (1990) principles of contrast and conventionality are two defining features of language, and to learn a language is to learn these two principles. For example, contrast has to do with learning words, because that is *what words do*: they mark contrasting states of affairs. Thus, it is not at all clear that the word-learning principles (constraints or biases) can exist prior to or independently of language learning. *As a description of word learning, however, constraints and principles cannot be taken for granted.*