LEARNING TO TALK ABOUT THE PROPERTIES OF OBJECTS: A NETWORK MODEL OF THE DEVELOPMENT OF DIMENSIONS

Linda B. Smith
Michael Gasser
Catherine M. Sandhofer

". . . the undeniable fact being that *any number of sensory sources, falling simultaneously on a mind* WHICH HAS NOT EXPERIENCED THEM SEPARATELY, *will fuse into a single undivided object for that mind.* William James, 1890, pg. 488, italics and capitalization as in the original).

A classic question in philosophy and psychology asks whether the object or its properties are more fundamental (see, e.g., Carnap, 1967; Locke, 1964; James, 1890; Boring, 1942). Clearly we perceive both—a dog is apprehended as an integral whole *and* as being big and brown and furry. But which perception is prior? This question is often interpreted in terms of the logical priority of parts and whole. And the contemporary concensus is that parts are logically and computationally prior; complex percepts and concepts are built from simpler primitives. More than 100 years ago, however, William James concluded that whole objects are experientially prior, that constituent properties are a secondary product of perceptual learning. This chapter provides support for James' conclusion.

Our starting point is the protracted course of children's acquisition of dimensional language, a lengthy process that includes the acquisition of dimensional terms and the development of selective attention. We explain

this developmental course by simulating it in a connectionist network. Our results suggest that dimensions are created, that they are the product of learning dimensional language.

I. The Developmental Data

Both Gentner and Rattermann (1991) and Smith (1989, 1993) provide extensive reviews of the development of dimensions. We highlight the main findings here.

A. Dimensional Terms Are Hard to Learn

Young children are rapid word learners, learning as many as nine words, per day (Carey, 1978). They are not, however, rapid learners of the words that refer to the perceptible properties of objects. Instead, the dimensional adjectives—words like *wet* and *soft* and *big* and *red*—are remarkably hard for young children to learn.

One line of evidence for this conclusion is the composition of early productive vocabularies; in these vocabularies, dimensional adjectives are rare or nonexistent. For example, in Stern and Stern's (1924) diary study of the acquisition of English, 78% of the words produced at 20 months were common nouns, the rest were verblike; *none* were adjectives. Similarly, in Nelson's (1973) study of 18 children learning English, fewer than 7% of the first 50 words were adjectives (see Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979). Dimensional adjectives are late for children learning other languages as well. In Dromi's (1987) study of one child learning Hebrew, only 4 of the first 337 words were adjectives. In a longitudinal study of the acquisition of Spanish by 328 children. Jackson-Maldonado, Thal, Marchman, and Bates (1993) found only one adjective among the 88 most common words. The finding that adjectives are infrequent in early vocabularies is notable given the frequency of common dimensional adjectives in adult language.

A second line of evidence consists of studies that attempt to teach children new words. Children as young as 18 months are one-trial learners of nouns (Woodward, Markman, & Fitzsimmons, 1994) and perhaps also verbs (Tomasello & Kruger, 1992). In contrast, studies attempting to teach young children adjective terms typically fail—even after as many as 2000 trials (Rice, 1980).

A final line of evidence showing the difficulty of learning dimension terms concerns children's errors. Long after children begin to use dimensional words, when they are as old as 3, 4, or even 5 years, their interpreta-

tions of dimensional terms are still errorful. This literature provides many examples of both within- and between-dimension errors, interpreting "big" to mean "tall" (Keil & Carroll, 1980), "little" to mean "big" (Donaldson & Wales, 1970), "big" to mean "bright" (Carey, 1982), "dark" to mean "loud" (Mark, Hammeal, Bornstein, 1987; Smith & Sera, 1992), and "red" to mean "green" (Binet, 1890).

In sum, children learn the names for object properties slowly and errorfully. We ask: Why is this learning so hard?

B. The Development of Selective Attention

The slow course of children's extraction of dimensions is evident in nonlanguage tasks as well. The evidence on children's difficulties are so substantial that the trend from wholistic to dimensionally differentiated perception has been offered as a principle of development (Gibson, 1966; Werner, 1957; Wohlwill, 1962). This large and well-documented literature shows that young children have difficulty in any task that requires them to attend selectively to one dimension. In these tasks, selective attention is measured by asking how much judgments along one dimension are disrupted when additional variation is added on another dimension. For example, a child is said to fail to attend selectively if when the child is asked to put red objects in one pile and blue objects in another, she makes more errors (or performs more slowly) when the red and blue things also vary in size and shape. Preschool children commonly fail to attend selectively in such tasks; indeed, they commonly fail to do them at all when irrelevant variation is added. In total, the evidence from discrimination learning tasks, classification tasks, matching-to-sample tasks, and same-different judgment tasks all show that young children perform poorly whenever they must attend to only one dimension of variation and ignore others (see Aslin & Smith, 1988; Smith, 1989; Gentner & Rattermann, 1991, for reviews). These results fit the idea that object properties fuse into an undivided whole in the minds of young children. Is this, then, why young children have difficulty learning dimensional terms?

C. Language Then Selective Attention

Children's difficulties in attending selectively might seem to suggest an explanation of children's difficulties in learning dimension words. If children cannot perceptually isolate properties from the object as a whole, then they cannot map dimensional adjectives onto object properties. But what drives the development of selective attention to dimension? The evidence suggests that learning dimensional language plays a key role.

Learning dimensional words and selective attention to dimensions are closely related achievements in children (Gentner & Rattermann, 1991; Smith, 1989). But overall, the evidence suggests that learning dimension words comes first. Although children learn attribute terms such as *big* and *wet* and *red* with difficulty, they often do so *years* before they can make judgments in nonlanguage tasks about one dimension *unaffected* by variation on other dimensions. Indeed, knowing the relevant dimensions words seem a prerequisite for successful selective attention in many tasks (Kendler, 1979; Ehri, 1976; Smith & Sera, 1992; see also Gentner & Rattermann, 1991, for a review). Other studies have shown that supplying children with words to describe properties and dimensions facilitates selective attention (Kotovsky & Gentner, 1996; Kendler, 1979). In sum, it appears that children first learn the names for attributes and *then* they become able to selectively attend to those properties (Sandhofer & Smith, 1996). In light of this evidence, Kotovsky and Gentner (1996) and Smith (1993) proposed that children's abstraction of perceptual dimensions is a consequence of learning dimension words. This chapter provides a demonstration of how learning dimensional language might lead to selective attention.

Fig. 1.   A representation of the idea that sensory, perceptual, and linguistic, representations are organized by the same dimensions.

II.   Toward a Model

To this end, our specific goal is a network that learns dimension words and through this learning develops the ability to selectively attend to dimensions. Three further issues influenced the architecture of the network and the learning tasks we presented to it.

A.   Different Dimensions at Different Levels

Figure 1 presents the classic view of dimensions: Different kinds of physical energy activate distinct sensory channels. These distinct sensory channels constitute the dimensions we can selectively attend to, perceive, and talk about. In brief, in the figure there are the same dimensions all the way down—from the dimensions we talk about to those we perceive to those that are given in the sensory system. By this view, the reason we *perceive* the dimensions we do is because these are the perceptual primitives our sensory system gives us.

The assumptions depicted in Fig. 1 are latent in much research in cognitive psychology: the search for primitive features (Treisman & Gelade, 1980), studies of dimensional crosstalk (Melara & Mounts, 1994), theories of category learning (Krushke, 1992) and in studies of infant perception and cognition (Husaim & Cohen, 1981; Coldren & Colombo, 1994).
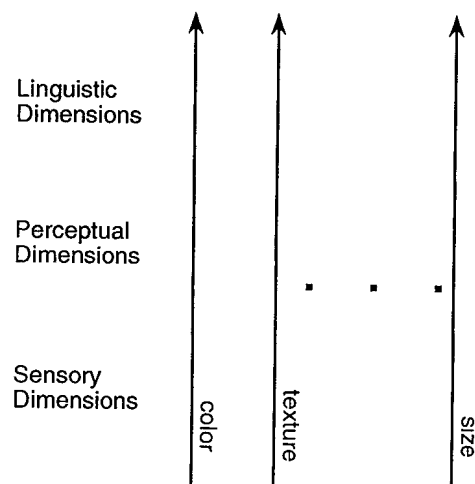
There are reasons to question this view, however. The first reason is the difficulty young children have in learning dimension words and in selectively attending to dimensions. The second reason is the fact that many of the dimensions that people talk about and use in their everyday lives to compare objects—the dimensions of size, shape, color, texture, wetness—do not correspond in a 1:1 fashion (and sometimes not all) to the features and dimensions identified at preperceptual or sensory levels of analyses (Treisman & Gelade, 1980; see also Hurvich & Jameson, 1957; Halff, Ortony, & Anderson, 1976). The third reason is the growing evidence that adults construct new features and dimensions as a consequence of learning categories (Freyd, 1983; Schyns, Goldstone, & Thibault, in press; Goldstone, 1995; Sanocki, 1991, 1992).

And, finally, is the evidence on the variability of the world's languages. The world's languages differ dramatically in how they lexicalize object properties. While all languages (apparently) have nouns and verbs as syntactic categories, some languages (Mandarin Chinese, Quechua) do not have a separate syntactic category of adjective. Rather, property words behave more like verbs. Thus there seems to be something less fundamental about how languages organize property terms. Second, languages slice up the sensory space in very different ways. For example, English "thick" brings together viscosity and width: Japanese "koi" brings together viscosity and

concentration (see Schachter, 1985). Differences among languages strongly suggest that dimensional words are not simple refections of fixed sensory attributes.

In light of this evidence of noncorresponding and changing dimensions at different levels in the cognitive system, we distinguish three senses of dimensions: (1) Sensory dimensions that structure the input to the cognitive system; (2) perceived dimensions that can be perceptually isolated, that is, attended to selectively; and (3) lexical dimensions that structure the relationships among words. Critically, our network will make no assumption that these three levels must correspond. Specifically, sensory dimensions will be hardwired into the network and linguistic dimensions will be given to the network in training. In simulations 1 and 2, linguistic and sensory dimensions will correspond and in simulation 3, they will not. In all cases, perceived dimensions must be constructed.

1. Three Kinds of Mappings

In learning dimension words, children learn three kinds of mappings. One mapping is between dimension words and object properties. For example, in learning the word "red" children associate "red" with red cups, red cars, red gum, red dresses, and so forth. These mappings are characteristically what one thinks of when one thinks of learning dimensional terms, and such mappings could be sufficient for the child to abstract the common property referred to by a single word. However, these word–property maps are not the only associations learned in the course of learning dimension words.

Children also learn word–word maps. The words "What color is it?" are associated with the words "red," "blue," and "green" but not "big." The word "size" is associated with the words "big" and "little"; "wet" is associated with "dry." There is evidence that children learn these associations and moreover that they sometimes do so before they learn the specific properties to which individual words refer (Cruse, 1977; Backsheider & Shatz, 1993; Clark, 1973; Carey, 1982). For example, Backsheider and Shatz showed that children who cannot correctly map a single color word to the right property often know to answer the question "What color is it?" only with color words.

Finally, *in the context of dimension words,* children also make property–property maps (Sandhofer & Smith, 1996). People put pairs and sets of objects before children and say such things as "These are both red," "They are the same color," and "These are the big ones." This learning context affords not just the mapping of words to properties but the mapping of properties to properties. Such simultaneous presentation of two objects with the linguistic specification of how they alike may encourage their comparison and the discovery of their common property. Gentner & Rattermann (1991); see also Kotovsky & Gentner, 1996) have suggested that such explicitly presented comparisons are crucial to the discovery of relations and kinds of similarity.

Accordingly, our network was designed to simultaneously learn word–property maps, word–word maps, and (word-)property–property maps. In simulation 1, we specifically examine the role of word–property maps and word–word maps in the creation of perceptually isolatable dimensions. In simulations 2 and 3, we examine the role of explicitly comparing objects on a linguistically specified dimension, that is the role of property–property maps.

2. Supervised Learning

Children learn dimension words through explicit teaching; adults provide both positive and negative evidence as to the properties to which dimension words refer (Callahan, 1990; Mervis, 1987; Snow, 1977; Wood, 1980). Adults as part of their regular interaction with young children ask such things as "What color is that?" "Is that big?" "Is kitty soft?". And, adults provide appropriate feedback. When a child labels a red object as *green* or a big one as *little,* parents tell them what the correct labels are. Thus back-propagation is a suitable learning algorithm because children's learning of dimensional terms is "supervised."

However, supervision as typically realized in connectionist networks is not perfectly appropriate as a model of children's word learning (see also Gasser & Smith, 1996). In traditional back-propogation, the connection weights on each learning trial are changed to increase the correct response and to decrease *all* other potential responses. This is like the parent saying to the child "This is red, not blue, not big, not wet, not soft, not bumpy. . . . Parents do not do this but instead explicitly reinforce correct answers ("yes, that's a red one") and provide negative feedback only when the child explicitly provides the wrong answer ("that's not red, its blue").

Traditional back-propagation is also inappropriate in the present case because the task of learning to label the multiple attributes of individual objects means that possible responses are not simply right or wrong. There are kinds and degrees of wrongness. Consider a big, black wet dog and the question "What color is it?" The answers *big* and *red* are both wrong. However, it seems unlikely that parents would respond to these errors in the same way. A toddler who answers the question "What color is it?" by correctly noting that the dog is *big* seems likely to hear a parental response of "yes, its a big dog, a big black dog."" A toddler who answers the same

question by saying *red* is likely to hear, instead, a parental response "its not red, its black."

Accordingly, we modified the back-propagation algorithm to fit these assumptions about the kinds of feedback provided by parents. The next section provides a detailed description of the network and the learning rule.

III. Network for Learning Dimensional Language

The architecture for our network instantiates three overlapping sets of mapping: property–word, word–word, and property–property. Figure 2 shows the network. Following convention each layer of units is represented as a box and the arrows between each layer represents complete connectivity between the layers of one unit and the next. There are three types of layers: (1) *Input layers*—these correspond to the sensory specification of an object and the linguistic context in which an object is perceived; (2) *Output layer* which corresponds to labels for attributes, words like *big, red, wet,* and *soft;* and (3) *Internal hidden layers.* We conceptualize the activity on these hidden layers as corresponding to representations at the level of conscious experience. We propose that the patterns of activations on these internal levels come, with the learning of language, to represent isolated attributes such that if the network selectively attends to the color
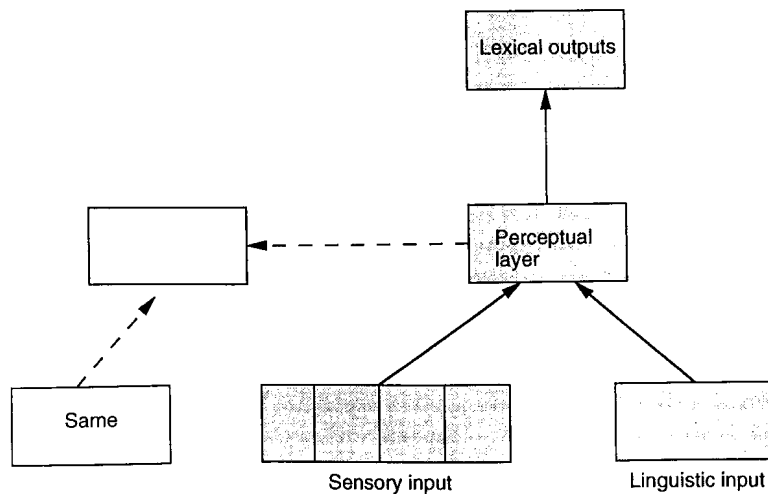
of a big black dog, blackness is isolated in the pattern of activity on the hidden layer.

We trained the network in two tasks common in children's learning of dimensional terms: (1) learning to label the attribute of a single object and (2) comparing objects. We describe the network in more detail by describing the layers and training procedure for each of these tasks separately.

A. LEARNING ATTRIBUTE NAMES: OBJECT–WORD AND WORD–WORD MAPS

The canonical test of a child's knowledge of dimensions terms—the one principally used by experimenters, parents, and educators—consists of presenting the child with an object and asking them a question about it: "What color is that?" "Is it big or little?" "What shape is it?" The child is thus presented with a single object along with linguistic input that specifies the relevant dimension and the child is asked to output the relevant dimensional term. Learning, to do this require mapping object properties to words, for example, the redness of the object to the word *red*. It also requires mapping dimensions words in the input to dimensions words in the output, for example, *what color* to *red*. One training task which we present to the network is based on this canonical test of dimension-word knowledge.

Four layers are involved; these are the shaded layers in Fig. 2. The sensory input and the linguistic-dimension input connect to a hidden layer which we call the perception layer. The activation of the perception layer in turn activates the units in the output layer. These output units represent words that name the properties of single objects along several dimensions. This portion of the network and its operation is the same as the network used by Gasser and Smith (1996) to model young children's faster learning of concrete nouns than dimensional adjectives.

1. Sensory Input

The object to be labeled is specified on independent sensory dimensions. In the present simulations, objects are specified on four (or in experiment 3 on five) sensory dimensions. The values of each sensory dimensions are represented using "thermometer" encoding (Harnad, Hanson, & Lubin, 1991). In this form of representation, units are activated in a series; the nth unit in the series is not activated until the activation on the n-1 unit is 1. Each dimension consists of 11 units, so two of the values along one 11-unit dimension are [1,1,1,1,1,1,.8,0,0,] or [1,1,1,.3,0,0,0,0,0,0,0].

2. Linguistic Input

The linguistic *input* signals the dimension specified in the question asked of the network. Each possibly relevant dimension is represented by a single

Fig. 2. The network trained to label attributes and to make comparisons. The shaded portion is the portion used to train attribute categories.

unit in this layer. Each of these units is thus associated with a class of possible answers—for example, activation of the linguistic input *color* is associated with the outputs *red, blue, green; size* is associated with the outputs *big, medium,* and *little.* Note, that there is no prior requirement that these linguistic dimensions match the sensory dimensions. In the first two simulations, there were three linguistic dimensions and thus three units in this layer.

Critically, from the perspective of the network, there is no distinction between the input activation that corresponds to the input object and that which corresponds to the linguistically specified dimension. From the network's point of view, there is just one input vector of 47 numbers jointly specifying an event in the world.

3. Perceptual Layer

Activations on the hidden perceptual layer represent the transient contents of immediate awareness. Thus these patterns of activation should change systematically with changes in the sensory and linguistic input such that the pattern of activation when labeling a big red dog as *red* will differ from that when labeling the same object as *big.* The patterns of activation that emerge on this layer in the context of different questions about different dimensions will thus constitute our definition of *perceived dimensions.* Importantly, the perceptual layer compresses the sensory dimensions so that the sensory dimensions are not directly recoverable in the patterns of activation on the perceptual layer. Rather, prior to learning, patterns on the perceptual layer are distributed, wholistic, representations of the input object. In this way, we embody James' claim in the opening quote that prior to learning, separate sensory sources "fuse into a single undivided object."

4. The Output Layer

The output layer consists of a single unit for each word, that is, each units corresponds to a dimensional adjective such as *red, green, big, little, rough,* and *smooth.* A +1 activation on the output unit represents the network's labeling of the input object with the corresponding word. A −1 activation represents the network's decision that the corresponding word is inappropriate for the input object, and a 0 activation represents an intermediate response, one that might be made if an object is described by a word that is not an appropriate answer to the question asked. For example, if asked "what color?" and given a large smooth red object, the output *red* would be represent by +1, the output *smooth* by 0, and the output *green* by −1. In this way we model parents likely nonresponse to true but not requested descriptors of the input object.

B. Making Comparisons: Property–Property Maps

In the course of learning dimension words, children are often presented with *sets* of objects and *told* how objects within the set are alike, for example, that two objects are "the same color" or "both soft." The portion of the network that learns attribute labels cannot learn from this sort of experience because it has no way of simultaneously representing more than one object and no way of comparing objects (see Smith, 1993). Our goal was to add to the network such that it could make internal comparisons of objects, the internal comparisons that would be triggered by being told that two objects are alike in some lexically specific way. To do this, we added two layers to the network used in learning attribute names: a Perceptual buffer had a "Same" unit (see Fig. 2). This "Same" unit simply turns on the comparison process which works in three steps:

1. Step 1. An object and a linguistic input specifying the relevant dimension are input to the network, activation is passed to the perceptual layer, and then the pattern on this layer is stored in the perceptual buffer.
2. Step 2. A second object is input along with the same linguistic input. This second object always has the same value as the first object on the linguistically specified dimension. Thus, at the end of this step, we have input two objects that are alike in some way and we have linguistically specified the dimension on which they are alike.
3. Step 3. The two patterns of activation on the perceptual buffer and the perceptual layer are compared: The pattern on the perceptual buffer is treated as a target for the pattern on the perceptual layer. The error is the difference between the two patterns. This error is back-propagated from the perceptual layer to the sensory and linguistic input layer. In this way the network is trained to make the internal representations for two objects which are the same on a linguistically specified dimension to be more similar to each other. In other words, the network is trained to find out what is alike about two objects "said" to be alike in some way.

C. Major Theoretical Claims

In summary, the network and training tasks instantiate four theoretical claims about how children learn to perceptually isolate and represent dimensions:

1. Prior to learning, perception is wholistic; separate sources of sensory information are combined into one unitary pattern.

2. Sensory, perceptual, and linguistic dimensions are distinct. Sensory dimensions are given in the biology and linguistic dimensions are in the language input; they need not correspond.
3. Perceptual dimensions are the product of learning dimensional language and thus may be constrained by both the sensory dimensions and the linguistic dimensions.
4. Two tasks are critical to the development of perceived dimensions: Learning to label the properties of objects which includes learning property–word maps and word–word maps and comparing objects on linguistically specified dimensions which creates property–property maps.

In the following simulation experiments, we demonstrate the plausibility of these ideas and show how they account for aspects of the developmental trend in children, yielding new insights and new predictions about the development of dimensions and selective attention.

Simulation 1 examines how learning to label attributes on linguistic dimensions that conform to sensory dimensions might lead to the perceptual isolation and representation of individual attributes. Simulation 2 examines the joint effects of learning to label attributes and of comparing objects on linguistically specified dimensions that conform to the sensory dimensions. Simulation 3 examines how this learning might create new perceived dimensions when linguistic dimensions do not conform to sensory dimensions.

IV.   Simulation 1: Learning Property Names

The central question behind this simulation is what one knows about dimensions when one has learned to call red things *red,* blue things *blue,* and big things *big.* Does labeling an attribute require the conscious isolation of that attribute from the other aspects of an object? This is an important developmental question because young children use dimensional terms *before* they can make other decisions about one dimension unaffected by variation on other dimensions, that is, before they can effectively selectively attend in nonlanguage tasks.

We addressed this question by asking the network to learn to label attributes on independent sensory dimensions. We then examined the character of the network's internal representations via what we call the Selective Attention Test. In addition, we assessed whether the network formed word–word maps on its way to acquiring attribute terms. This simulation involves only learning to label the properties of objects presented one at a time and thus only the shaded layers in Fig. 2.

A.   METHOD

1.   Training

The network was taught to answer three questions—each a request for the name of an attribute on one sensory dimension. In total, the network was taught nine attribute categories. Each adjective category was organized by a range of values on one sensory dimension; specifically, each attribute label referred to .33 of the maximum possible range on *one* sensory dimension. Three attribute categories each were defined for three of the four sensory dimensions. There were no terms that corresponded to values on the fourth sensory dimension.

Each trial consisted of the presentation of an object on the sensory layer and a linguistic input. There were three possible linguistic inputs; what might be characterized as the questions *What color is it?, What size is it?,* and *What texture is it?* The first linguistic input was associated with ranges of variation on the first sensory dimension, the second with ranges of variation on the second sensory dimension, and the third with ranges of variation on the third sensory dimension. For each trial, a linguistic input was randomly selected and then an object was randomly generated such that each attribute term was the correct answer equally often. The network was trained on 15,000 randomly generated inputs and tested every 2500 inputs.

2.   Selective Attention Test

We tested the network's ability to isolate attributes every 2500 trials by examining the patterns of activation on the perceptual layer for pairs of objects that were either same or different on the linguistically speicfied dimension. The idea is this: If the network has abstracted the property *red* from all other properties, then the pattern of activation on the perceptual layer should be the same when the network is asked "What color?" and given and a big, red, bumpy, rounded object *and* when it is asked "What color?" and given a little, red, smooth, angular object.

For the Selective Attention Test, we specifically examined two kinds of pairs: (1) Same-on-Relevant-Dimension pairs consisted of two inputs that were the same on the linguistically specified sensory dimension but different on the other three dimensions; and (2) Different-on-Relevant-Dimension pairs consisted of two inputs that were different on the linguistically specified dimension but the same on the other three. On each Selective Attention Test trial, each member of the pair was input and its resulting pattern of activation on the perceptual layer was stored. Then the second member of the pair was input and its resulting pattern of activation on the perceptual

layer was stored. The dependent measure was the Euclidean distance between these two patterns of activation.

Prior to learning, the Same-on-Relevant Dimension items should yield *more dissimilar* patterns of activation on the hidden layer than the Different-on-Relevant Dimension items because the Same pairs are alike on one sensory dimension but different on three and thus are more wholistically different than are the Different pairs which are different on one sensory dimension but alike on three. If, however, in learning the attribute terms, the network learns to perceptually isolate and selectively attend to the linguistically specified dimension, then the patterns of activation for the Same-on-Relevant dimension items should become more similar, less distant, than the Different-on-Relevant dimension items.

This experiment (and the others that follow) was conducted ten times with the network starting learning each time with different initial and randomly determined connection weights. Each of the 10 runs also employed different randomly generated inputs. The results are reported as means over the 10 runs. We considered two definitions of a correct response: (1) the network was correct if the most highly activated output was correct or (2) the network was correct if the activation of the correct output unit was above a predetermined threshold. Both measures lead to the same pattern of results and conclusions. We report the results in all of the experiments in terms of the output with the highest activation.

3. Results and Discussion

Figure 3 shows the network's proportion correct labeling of attributes on the linguistically specified dimension as a function of training. By the 5000th input, the network was nearly always correct; given an object and a linguistic input that specified the relevant dimension, the network correctly labeled the appropriate attribute.

Figure 4 shows the results of the Selective Attention Test: The Euclidean distance between patterns of activation on the perceptual layer as a function of training for the Same-on-Relevant dimension pairs and the Different-on-Relevant dimensions pairs. At the beginning of training, the pattern of activations for the Different pairs were less distant, more similar, than the Same pairs which is to be expected because the Different items shared values on three sensory dimensions but the Same items shared a value on only one sensory dimension. With training, the distance between these Different pairs increases but the distance between the Same pairs does not decrease. That is, the network *did not* learn to perceptually isolate a common sensory property by learning to label it. This failure to isolate properties stands in marked contast to the high performance of the network in
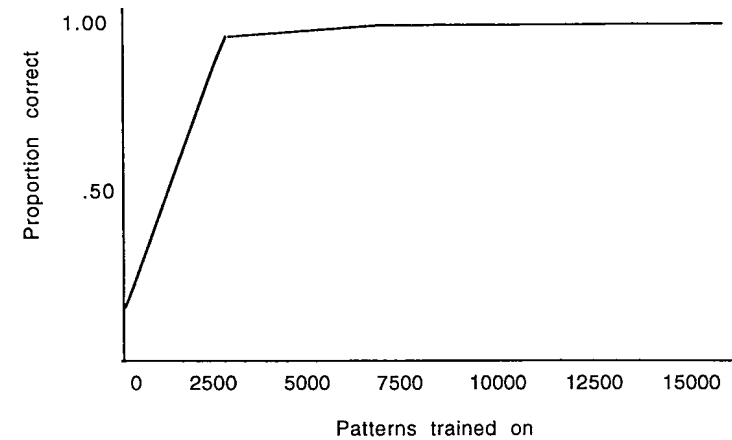

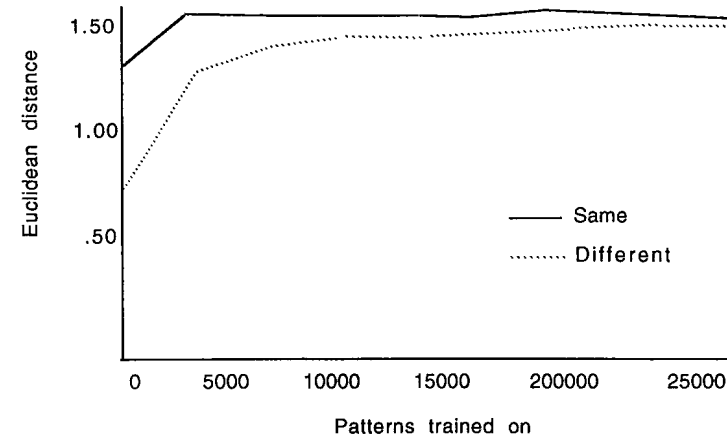Fig. 3.   Proportion correct responses in simulation 1 as a function of the number of training patterns.

Fig. 4.   The Euclidean distance between patterns of activation on the hidden later for two input objects that are the same on the queried dimension and different on all other dimensions (*solid line*) and for two input objects that are different on the queried dimension and the same on the other three (*dashed line*) as a function of training patterns.

learning to label attributes. The network learned to "call" blue objects *blue,* red objects *red,* big objects *big,* bumpy objects *bumpy,* and so on with accuracy but without learning to isolate the common sensory values that define these attribute categories.

The finding that correctly labeling attributes does not require the perceptual isolation of the attributes is not surprising in one sense. To succeed, the network only needs to find partitions of the hidden-layer activation space such that all instances of a dimensional term fall on one side of a boundary. The task does not demand that the network find an identical pattern of activation, an invariant, for all red things; even though there is one available in the sensory input.

However, the present results *are* surprising in the context of typical inferences in developmental psychology which take verbal behavior as a close indicator of underlying concepts and their constituent structure (e.g., Gelman & Coley, 1991; Keil, 1989). In this context, labeling red things *red* is prima facie evidence for "having the concept red." However, the present results show that "having the concept" need not mean having abstracted a common component. Notice that the behavior of the network does fit the behavior of children who use many dimensional terms correctly before they effectively attend selectively to the properties those terms name. The clear implication is that young children can correctly use a dimension word *before* they attend selectively to the property named by that word.

Although the network did not learn to isolate perceptually the attribute defining a lexical category, it did learn something about linguistic dimensions, forming robust word–word maps between linguistic inputs and outputs, and like young children it did so prior to learning to correctly label properties. Figure 5 shows the proportion of errors that were *within-dimension* errors. Given three dimensions each with three attribute categories, the proportion of within-dimension errors expected by chance is .25. As is apparent, these errors increase with training and after training occur more often than expected by chance. Again, this behavior is like that of young children who know the class of appropriate answers to a question before they correctly use individual dimensional terms.

In sum, the network learned to label properties without discovering the invariant "sensory" properties in the input. The network could succeed in this way because to be right it only had to find a set of connection weights that worked well enough. The results suggest that dimensional adjectives may be initially learned by children as broad multidimensional categories even when they correspond to "given" and separate sensory primitives.
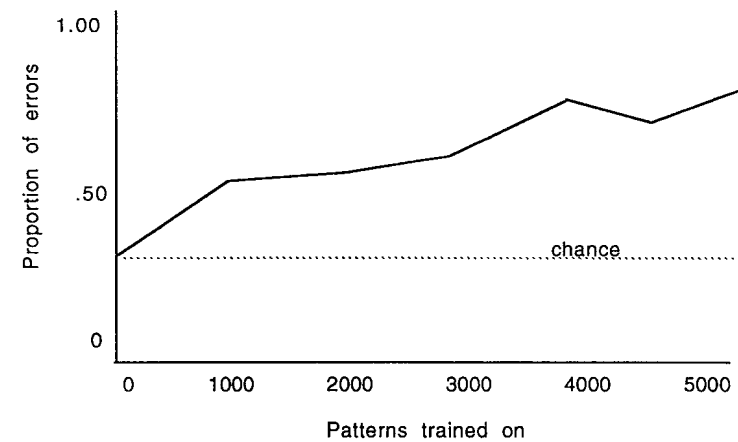

Fig. 5. Proportion of errors that were within-dimension errors as a function training patterns in simulation 1.

V. Simulation 2: Learning to Selectively Attend

If perceived dimensions are not "givens" but must be discovered, and if they are not necessarily discovered by learning to label the attributes of objects, how are they learned? In this simulation, we ask if comparison of objects on a linguistically specified dimension results in the perceptual isolation of individual dimensions. That is, in this simulation, the network is asked to learn three sets of maps: word–word maps between linguistic inputs and outputs, property–word maps between input objects and linguistic outputs, and property–property maps between objects that are specified to be alike on a particular dimension. Learning this third relational map should teach the network to find common properties and to filter out irrelevant information from linguistically unspecified dimensions. The issue is how this training interacts with the learning of attribute categories.

We examined two training procedures. In the Joint-learning condition, we taught the network to label attributes and we also presented it with pairs of objects to compare on the linguistically specified dimension. Both kinds of training trials alternated from the start of training. In the Attributes-then-Comparison condition, we taught the network to label attributes as in simulation 1 and then after this first task was mastered we introduced the comparison task.

A. METHOD

1. Stimuli

The stimuli for learning to label attributes were identical to those used in simulation 1. The stimuli for comparison training were pairs of inputs. The two input patterns had identical values on the sensory dimension specified by the linguistic input and differed by at least 30% of the range on each of the other three dimensions. On each comparison training trial, pairs of inputs were randomly generated to meet these constraints. The stimuli for the Selective Attention Test were identical to those used in simulation 1.

2. Procedure

Training to label attributes was conducted as in simulation 1. For each comparison training trial, a lexical dimension was first picked randomly to be the dimension specified by the linguistic input. Next, a pair of objects identical on the linguistically specified sensory dimension and different on the remaining three was generated. The first item in the pair was presented on the sensory layer, together with the appropriate linguistic input. The resulting pattern on the perceptual layer was then saved on the perceptual buffer. Next, the second object was presented on the sensory layer along with the appropriate linguistic input. The resulting (second) pattern on the perceptual layer was then compared to the (first) pattern on the buffer. The pattern on the buffer served as the target.

In the Joint-learning condition, comparison training was introduced along with training to label attributes from the start of training. The two tasks alternated during training.

For the Attributes-then-comparison condition, the network was trained in the attributes task alone for the first 5000 inputs. At 5000 inputs, comparison training was introduced and the two tasks alternated for the remainder of training.

In both training regimens, training continued until 15,000 patterns (7500 for each task) had been presented. Following every 2500 inputs, the network was tested in the Selective Attention Test also as in experiment 1.

3. Results and Discussion

Figure 6 shows the network's performance labeling attributes in the Joint training condition, when the network learned to do both tasks from the start of training. As shown in the figure, the network learned the attribute names in this condition more slowly than in simulation 1 when it was trained only in the attribute labeling task. Learning to label attributes is clearly
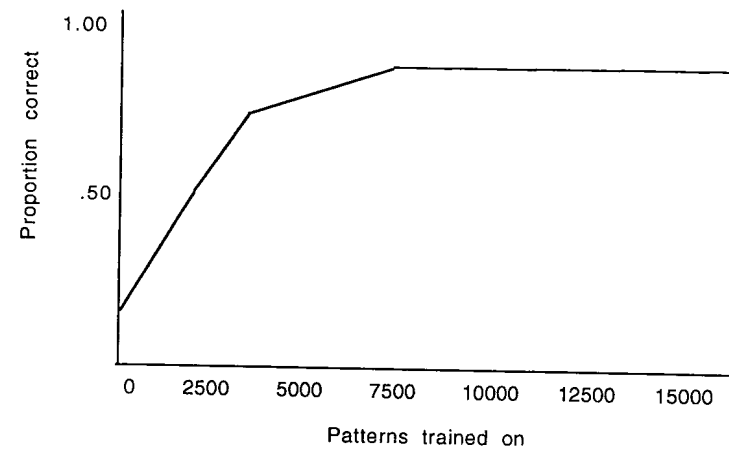
Fig. 6.   Proportion correct responses in the attribute labelling task under the Joint training regimen in simulation 2.

made more difficult by simulataneously learning to compare objects along the same lexical dimensions.

However, the results in the Selective Attention task, shown in Fig. 7, show that this joint training resulted in different knowledge about attributes
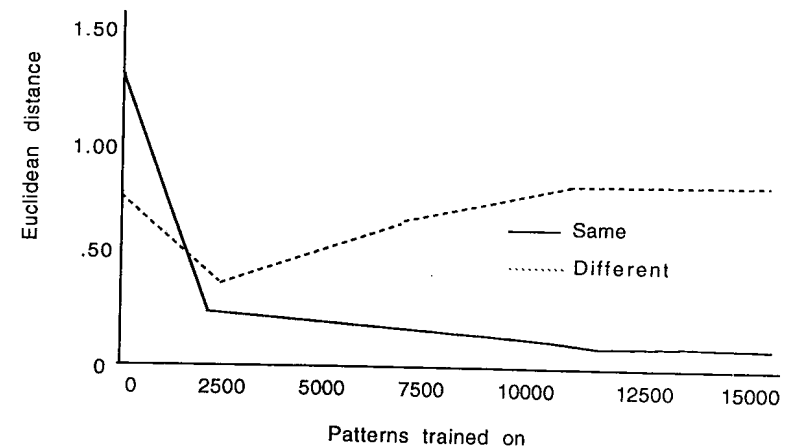

Fig. 7.   The Euclidean distance between patterns of activation on the hidden layer for pairs of inputs same on the queried dimension (solid line) or different on the queried dimension (dashed line) as a function of training patterns in the Joint training regimen of simulation 2.

than had the training in simulation 1. At the start of learning, the distance (dissimilarity) of patterns of activation on the perceptual layer of Same-on-Relevant dimension pairs is greater than Different-on-Relevant dimension pairs. This is to be expected since Same pairs are alike only on the lexically specified dimension but different on the remaining three, whereas Different pairs are different on the lexically specified dimension but same on the other three. However, briefly into Joint training, the pattern reverses so that the network now internally represents objects that are alike on a lexically specified dimension as being more alike than objects that differ on the lexically specified dimension. The network has learned to isolate on the lexically specified dimension. The network has learned to isolate the property common to objects labeled by the same term.

Figure 8 shows how well the network learned to label attributes when it was first trained to name object properties and then given joint training on both naming attributes and comparing objects. This training regimen is clearly best for learning to label attributes; learning is rapid and largely unaffected by the introduction of comparison training. Thus, although comparing objects on single dimensions slows learning to label attributes when they are both taught from the beginning, the *addition* of this task does not disrupt the labeling of already learned attribute names. This is intriguing because what the network knows about attributes is different before and after comparison training.

The results of the Selective Attention Test in the Attribute-then-Comparison condition are shown in Fig. 9. When the comparison task is introduced, the network quickly learns to selectively attend: patterns of activation for inputs that are the same on the lexically specified dimension become highly similar and patterns of activation for inputs that differ on the lexically specified dimension become very different. By 10,000 training patterns, performance in this condition on the Selective Attention Tasks equals that of the network on the Selective Attention Task in the Joint Training condition. Overall, this incremental learning schedule, by which the network learns to solve one task before starting on another, is superior to the one in which the network is faced with both tasks through out training. This fact and the fact that the addition of comparison training does not disrupt labeling object properties suggests that when learning attribute terms, the network learns attribute representations that *are in the right direction,* representations that are then refined by comparison training.

The results show that the network can learn to attend selectively and isolate sensory dimensions but that it needs explicit comparison training to do so. Learning to label attributes is not sufficient for the perceptual isolation of properties. Importantly, comparison training as realized in this network is not sufficient either. The perceptual isolation of attributes cannot be learned by learning property–property maps alone because the use of one input as a target for the other (i.e., the task of making the two patterns
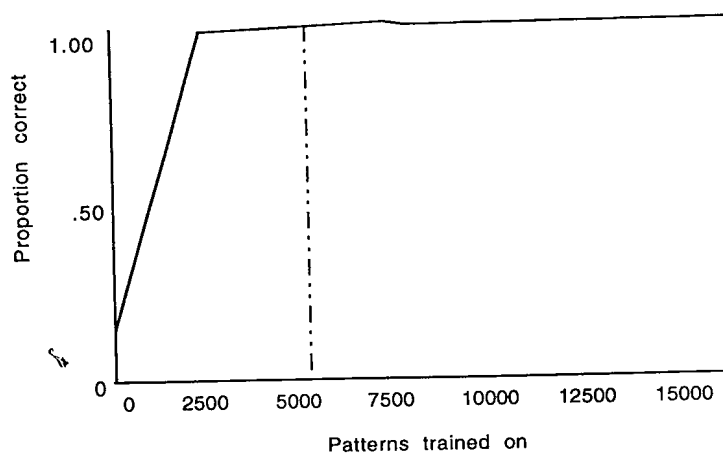

Fig. 8.   Proportion correct responses in the attribute labeling condition under the Incremental training regimen. The network receives categorization training only for the first 5000 trials and then receives alternating categorization and comparison training for the remainder.
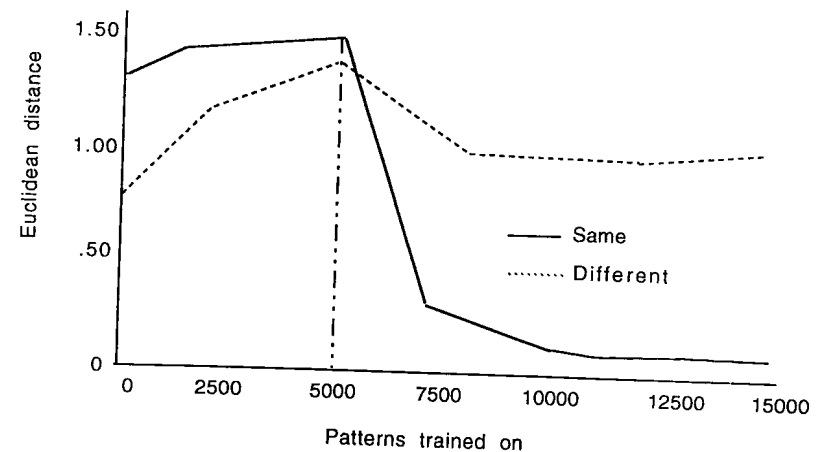

Fig. 9.   The Euclidean distance between patterns of activation on the hidden layer for pairs of inputs same on the queried dimension (*solid line*) or different on the queried dimension (*dashed line*) as a function of training patterns in the Incremental training regimen of simulation 2. Comparison training introduced after 5000 trails.

of activation on the comparison buffer and perceptual layer the same)
drives *all* patterns of activation on these two layers to become alike. In a
preliminary simulation in which we trained the network only in the compari-
son of objects, we found that after 2500 input patterns, the Euclidean
distance between all inputs (whether the same or different on the relevant
dimension) was lessn than .03. Thus, training to label attributes alone is
insufficient *and* training in comparison alone is insufficient. Success requires
the formation of multiple mappings through the same connection weights;
training in word–property maps, word–word maps, and property–property
maps. Apparently, learning attribute labels keeps the patterns of activations
on the perceptual layer sufficiently different for lexically different attributes
and comparison training along linguistically specified dimensions enables
the network to isolate the common property.

In sum, the network learned to attend selectively to sensory dimensions
by learning about linguistic dimensions; that is, by learning to name attri-
butes and by finding what is alike about objects that are said to be the
same on a nameable dimension. In the final simulation, we ask whether
the network can learn to isolate attributes on dimensions that are not
directly given in the sensory input.

VI. Simulation 3: Creating A New Dimension

In the previous two experiments, the network was asked to discover sensory
dimensions—to abstract what was common in the input description of
individual objects. The network discovered sensory dimensions and learned
to attend selectively to them by learning linguistic dimensions. In a sense,
the network learned to "map" linguistic dimensions onto sensory dimen-
sions. This was accomplished, however, not by gaining direct access to the
sensory dimensions which are compressed at the perceptual level but by
reconstructing them. Thus the question of this experiment: If the network
can reconstruct sensory dimensions, can it also discover and perceptually
isolate dimensions that are not simple sensory dimensions? This is important
question because people appear to be able to learn new dimensions (Thi-
bault & Schyns, 1995; Goldstone, 1995), and to talk about dimensions
that do not directly correspond to the features and dimensions that are
independent at early stages of sensory processing.

Accordingly, in this final simulation, we taught the network to label
attributes and make comparisons on a dimension that did not correspond
to any sensory dimension.

A. Method

The network was the same as that used in simulations 1 and 2 except for
two changes: (a) objects were specified along five sensory dimensions, and

(b) the linguistic input specified four linguistic dimensions. Each sensory
dimension was organized as in simulations 1 and 2. The network was taught
12 attribute labels. Three labels referred to values on sensory dimension
1 and were associated with linguistic dimension 1; three attribute labels
referred to values on sensory dimension 2 and were associated with linguistic
dimension 2, and two attribute labels referred to values on sensory dimen-
sion 3 and were associated with linguistic dimension 3. We call the poten-
tially isolatable dimensions that correspond to the sensory dimensions,
Simple Dimensions. The range of values on each of these sensory dimen-
sions named by each attribute term was 20% of the dimension. Thus, some
portion of possible values on each dimension were not specifically labeled.

Four of the attribute labels, in contrast, were defined in terms of values
on sensory dimensions 4 and 5 and were associated with a single linguistic
dimension, linguistic dimension 5. Each of the attribute terms on this Com-
plex Dimension was defined as a point along and near a linear relation
between sensory dimensions 4 and 5. Specifically, values on the complex
dimension were constrained as follows:

$$.8 < \text{sensory dimension 4} + \text{sensory dimension 5} < 1.0$$

and the four attributes were defined as ranges of values within these con-
traints. In the possible space of all visual inputs along the five sensory
dimensions, this inequality defines a complex dimension with a rectangular
shape. Each of the four complex attributes refers to a subregion that is 20%
of the complex dimension. That is, there were boundary regions between
attributes that were not labeled.

The training procedure was identical to that used in the Attributes-then-
Comparison condition of simulation 2.

2. *Results and Discussion*

Figure 10 shows the performance of the network in the task of labeling
attributes: the four dimensions as a function of training trials. The dotted
vertical line indicates the point at which comparison training was intro-
duced; that is, the point at which the network was asked to discover what
the same about two objects that were labeled as same on a specific dimen-
sion. As is apparent, prior to the introduction of the comparison task, the
network readily learned to label *all* the attributes—those on the Simple
and those on the Complex dimension.

As is also apparent in Fig. 10, performance declines somewhat when the
comparison training is introduced as when the network is asked to discover
what is the same about two objects that are the *same on a dimension,* but
as in Simulation 2, it returns to near perfect performance. Performance on

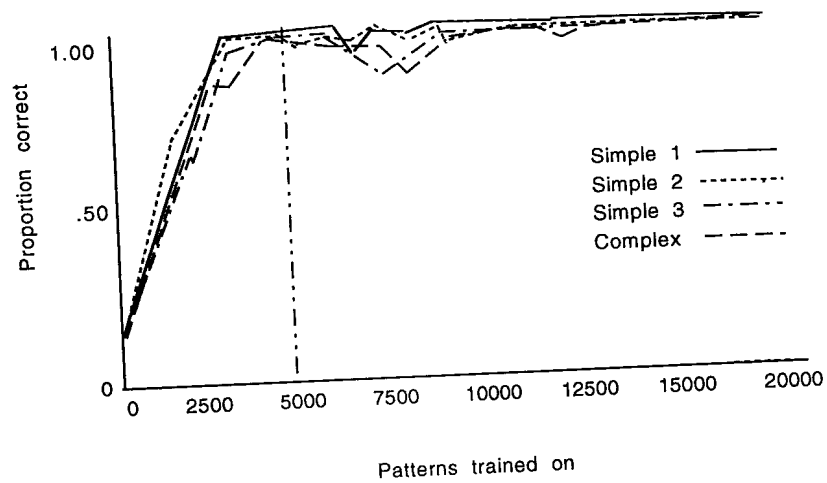Linda B. Smith, Michael Gasser, and Catherine M. Sandhofer

Fig. 10.   Proportion correct responses in the attribute labeling condition under the Incremental training regimen for the three Simple dimensions that correspond to sensory dimensions and for the Complex dimension. Comparison training was introduced after the first 5000 trials.

the complex dimension appears no different at this point than performance on the simple dimensions.

Figure 11 shows performance in the selective attention task for each dimension. The dependent measure is the distance between patterns of activation on the perceptual layer for visual inputs that are the same on the relevant dimension (and different on the others) or are different on the relevant dimension and the same on all others. Notice first the performance prior to the introduction of Comparison training. At this point, the network accurately labels attributes on all four dimensions. But for none of them does this correct labeling mean the perceptual isolation of the labeled attribute. *And,* for the complex dimension, the correct labeling occurs without a discretely localized representation of the dimension or its attributes *anywhere* in the network. This clearly shows how the outward behavior of a system may be a very poor guide to the internal processes that make that behavior.

However, the network did learn to attend selectively—to isolate what is common—given comparison training; and it did so for all four dimensions. Again, performance on the complex dimension is as good as performance on the best simple dimensions. The patterns of activation at the perceptual layer for objects that share a value on the relevant dimension are, at the end of training, considerably more similar than the patterns of activation for inputs that are different on the relevant dimension.
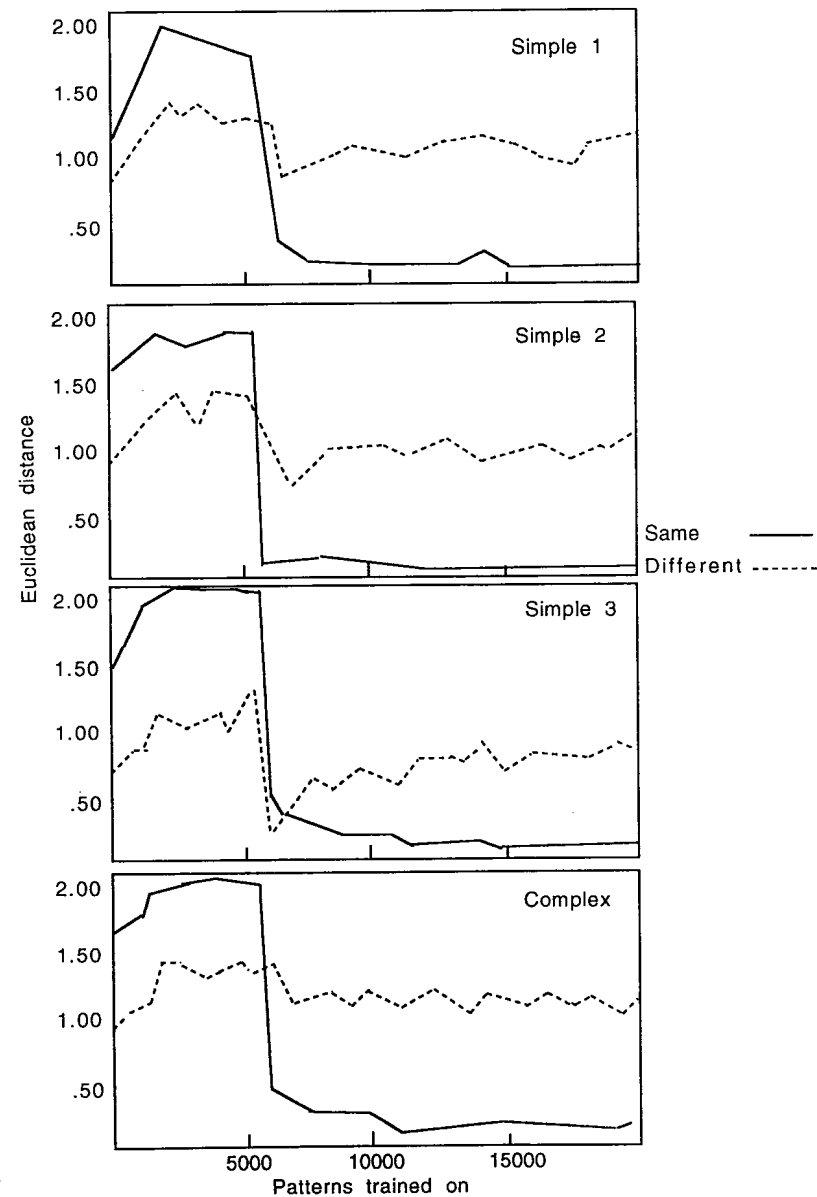

Fig. 11.   The Euclidean distance between patterns of activation on the hidden layer for pairs of inputs same on the queried dimension (*solid line*) or different on the queried dimension (*dashed line*) as a function of training patterns in the Incremental training regimen of simulation for the three Simple and one Complex dimensions. Comparison training introduced after 5000 trials.

These results show that the network can learn to name attributes and attend selectively to a dimension that does not correspond to a single sensory dimensions as well as it can learn about sensory dimensions. Isolatable dimensions can be created that have no a priori existence in the system.

A final aspect of these results, like those in simulation 2, points again to how the outward behavior of a complex network need not reflect the same underlying processes. Prior to training in the explicit comparison task, the network outputs attributes labels accurately and often comparison training, the network still outputs attribute labels correctly, but the network's similar outward behavior in the two cases is based on radically different "representations." Moreover, the shift from one form of representation to another is only discernible in what are quite small and transient disruptions in the labeling of attributes. Further, in this experiment, the network learned to label attributes on simple and complex dimensions in comparable ways with nothing in its patterns of outputs to indicate that some attribute terms label values on the sensory "primitives," whereas others label values on complex learned dimensions.

VII.   Developmental Implications

This chapter began with the developmental evidence on children's slow acquisition of dimensional terms and their difficulties in selective attention tasks. Our model provides new insights by showing how dimensions may exist at multiple noncorresponding levels, by showing how language learning involves learning a system of mappings between these levels, and by showing how jointly learning these mappings may create perceptual dimensions.

A.   The Starting Point Is Wholistic Perception

The starting point for our developmental account is the wholistic compression of the sensory dimensions by the hidden layer, the layer we take as corresponding to subjective experience. The theoretical idea is that learning dimensional adjectives is hard because the system is structured to perceive objects as unitary wholes. This aspect of our model fits Markman's (1989) proposal about a "whole-object" bias in early word learning. She explains children's special difficulties in learning dimensional adjectives in terms of constraints on their initial hypotheses about possible word meanings. She proposed that young children first learning words assume that those words refer to individual whole objects rather than to the component properties of objects. This is a learning principle that promotes the acquisition of nouns because objects in the same nominal category are typically similar

across many interrelated and correlated properties (e.g., Markman, 1989; Rosch & Mervis, 1975). Our network with its compression of sensory dimensions provides a mechanistic implementation of the whole-object assumption. In other works, we have shown that this network, like young children, is biased to learn noun meanings and that it learns nominal categories organized across many dimensions more rapidly than it learns dimensional-adjective categories (Gasser & Smith, in press).

What the present model adds to Markman's original idea is the further proposal that children must actually construct perceptual dimensions and that they do so through learning language. The developmental evidence has suggested this possibility to others (Bruner, 1957; Kendler, 1979; Gentner & Rattermann, 1991; Kotovsky & Gentner, 1996; Smith & Kemler, 1978; Smith, 1984; Smith, 1989). The present results demonstrate that such learning could occur.

Our network with its wholistic compression of sensory dimensions contrasts with a well-known model of dimensional learning in the adult literature—Kruschke's (1992) ALCOVE model. ALCOVE retains the separateness of distinct sensory dimensions across layers in the network by utilizing dimensionally distinct learning weights. ALCOVE thus instantiates the ideas depicted in Fig. 1: the same dimensions from the bottom of the system to the top. ALCOVE has not been applied to developmental phenomena but learning by this network fits well the pattern of adult learning in many simple categorization tasks (e.g., learning to classify a small set of instances into two mutually exclusive groups). In these tasks, ALCOVE, like adults, rapidly learns categories organized by one dimension. In brief, ALCOVE models well the end-state structure, how category learning works after separate dimensions are formed.

The question is whether ALCOVE can be made to fit the developmental data? One possible way is make changes in dimension weights as a function of training more sluggish such that the network continues to distribute attention across all dimensions longer through the learning process. Without actual simulations, it is difficult to know how much of the developmental evidence could be adequately captured by this approach. However, even if such an approach did work, ALCOVE would still offer no explanation of the noncorrespondence of the dimensions we talk about and those defined by sensory psychologists, no explanation of perceptual learning that creates new dimensions and properties, and no explanation of the variability in dimensional terms across languages.

B.   Separable Dimensions and Selective Attention

The classic definition of a psychological dimension is context independency: judgments of one dimension should be independent of values on other

dimension (e.g., Boring, 1933). Thus, Garner (1974) defined experientially separable dimensions as those that afforded (near) perfect selective attention to one dimension at time. This is the same benchmark we required of our network. This benchmark was achieved late in learning by the network and required more by way of training than merely learning categories well organized by one dimension. Dimensional separability by Garner's definitions also develops quite late in children (e.g., Smith & Kemler, 1978). Thus, again, the course of developments by the model mimics that observed in children.

One could argue, however, that the benchmark of perfect selective attention is too high. By this view, the network and children could be said to attend selectively to and "have" dimensions earlier; the only problem is that they selectively attend imperfectly. In order to label attribute categories, the network must have learned something about the attributes. It must have formed a partition of patterns of activation such that patterns turning on different outputs were on different sides of the partition. The problem is what evidence should count as indicating psychological dimensions. If the criteria for defining a dimension is independence from other dimensions, then imperfect selective attention would seem the product of partially created dimensions.

C. Different Dimensions at Different Levels

A central claim of in our account is that dimensions exist at multiple levels: there are, initially, sensory dimensions that code the physical input and linguistic dimensions that organize dimensional adjectives. Perceptual dimensions are constructed, we propose, by learning multiple mappings between and within these levels. This idea of different dimensions at different levels helps explain some peculiarities in the developmental literature. The most important is the evidence purporting to show that infants can do what young children cannot: selectively attend to dimensions.

The evidence for selective attention to dimensions in infants derives from habituation studies (e.g., Cohen & Oakes, 1993; Fantz, 1963; Bornstein, 1985; Coldren & Colombo, 1994). In these studies, infants are shown repeated examples of objects that are alike on some dimension but vary on a second dimension. For example, as illustrated in Fig. 12, infants might be repeatedly shown red squares, red circles, and red triangles until looking at these objects declines. After habituation, the infant is presented with test trials. For example, infants might be shown an object that differed on the previously constant (relevant) dimension but which matches one of the exemplars on another (irrelevant) dimension: e.g., a green square. Or infants might be shown at this point an object that matches the exemplars on the
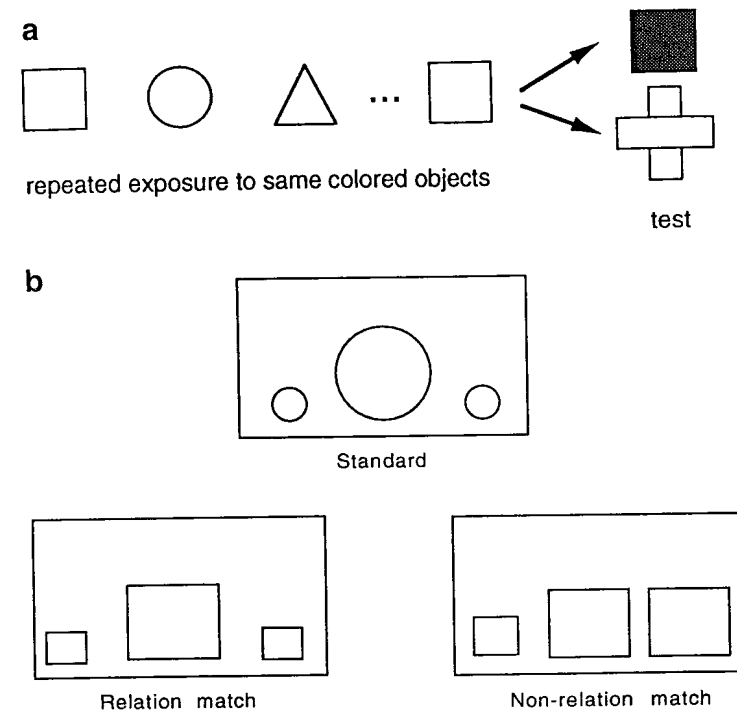

Fig. 12.   (A) A series of possible stimuli in a habituation study. (B) Illustration of a possible stimulus set in Kotovsky and Gentner's experiment.

relevant dimension but differs from them on another dimension, e.g., a red cross. An increase in looking time to a test object suggests discrimination of that object from the habituation exemplers. The standard result is that infants look longer at test objects that differ on the relevant dimension (i.e., the green square in the example) than ones that differ on an irrelevant dimension (i.e., the red cross in the example). These results suggest selective attention to the relevant dimension (color) and the ignoring of irrelevant properties.

Two kinds of accounts have been offered of young children's difficulties in the face of infants successful selective attention. Several investigators (Smith, 1989; Aslin & Smith, 1988; Coldren & Colombo, 1994) take the data at face value and conclude that because infants can attend selectively to dimensions, then there must be a fixed and universal set of dimensions that structure perceptual experience. In this view, preschool children fail to attend selectively for other reasons. Noting the close relationship between

the acquisition of dimensional language and successful selective attention. Coldren and Colombo went so far as to suggest that preschool children *lose* their ability to attend selectively because they are initially confused by dimensional language. This and the various other explanations of children's failure to attend selectively in terms of a failure to selectively attend are all unsatisfying in their circularity.

Kemler (1982), in contrast, argued that the infant data do not show what they seem to show. She correctly pointed out that the infant data do not conclusively demonstrate *perfect* selective attention to dimensions, which is the benchmark of dimensional separability in adults and the benchmark that young children fail to meet. Kemler argued that the same pattern of habituation-dishabituation could emerge given psychologically arbitrary (that is made-up) dimensions. Her point is that selective dishabituation does not guarantee preexisting internally represented dimensions. This idea is supported by our findings that the selective naming of red things *red* does not guarantee the abstraction of redness from other properties.

Our model also suggests a third account. Infants and preschool children's "dimensional" judgments may not conform because they do tap into dimensions at the same level. Habituation studies may measure sensory dimensions, whereas speeded classification and comparison tasks may measure perceptual dimensions. There may be no single sense of dimension that applies in all tasks and from the bottom to the top of the cognitive system.

Clearly, what is needed is a unified mechanistic account of selective habituation to dimensions in infancy, the learning of dimensional words, and the development of selective attention. These present simulations make a start in this direction.

D.   Dimensions Are Made Through Multiple Mappings

The network learned mappings between questions asked about objects and classes of possible answers, it learned mappings between words and object properties, and it learned mappings between the internal representations of pairs of objects and a linguistic description of their relatedness. *All* of these mappings appear essential to the construction of perceptual dimensions, but none are sufficient. The task is not solvable without word–word maps from inputs to outputs. Without such maps, the network cannot know, given a big red smooth thing, whether it is to attend to red or smooth or big. Isolated perceptual dimensions cannot be formed from object-word maps. Rather, perceptual dimensions require comparison and explicit learning of property–property maps. Perceptual dimensions, however, cannot be formed by comparison alone. Gains from comparison must be bootstrapped on the prior learning of attributes categories.

This bootstrapping relation between comparison and category learning is reminiscent of Karmiloff-Smith's (1986) ideas about "re-representation." She proposed that earlier representations, often with the help of langauge, are compared to one another and re-represented in more abstract and context-free form. Our network offers one instantiation of how this might happen: The explicit comparison of two objects both called *red* encourages the re-presentation of redness as a discrete entity divorced from size or texture or other aspects of the whole object.

The importance of the comparison process has also been argued by Gentner and colleagues (Gentner & Rattermann, 1991; Kotovsky & Gentner, 1996). They also report relevant evidence. In one study, Kotovsky and Gentner examined children's ability to match patterns such as those shown in Fig. 12. Given the standard shown at the top, the child's task was to choose which of the bottom two patterns "was like" the top ones. This is a very hard task for preschool children. Kotovsky and Gentner found that children's performances were helped by teaching them labels for the relations. For example, the experimenter pointed to the matching endpoints and told the children that the exemplar won because it was "even." Children were trained on a set in which only the size varied and then successfully transfered to sets that demanded more difficult pattern matches. Kotovsky and Gentner concluded that repeated comparisons of objects said to be "even," enabled children to abstract the relevant relation.

E.   Language Helps Make Dimensions

In the simulations presented here, langauge is a strong force shaping perceptual dimensions. Is the implication, then, that language is necessary for the perceptual isolation of dimensions? Language is not logically necessary; the simulations show that. A learning task in which the inputs and outputs were not words but other contexts and other responses could accomplish the same set of mappings.

Experiments with nonhuman animals also suggest that the key is the system of mappings and not language per se. Many experiments in this literature demonstrate the acquisition and generalization of attribute categories—categories demanding the same response to objects alike in a particular way (e.g., Rescorla & Wagner, 1972; MacIntosh, 1965). There are fewer experiments that clearly demonstrate comparison training (see Premack, 1976, for a discussion). One comparison task, however, that has been widely used is matching-to-sample. In this task, the animal is presented with three objects and must find the two objects in a set of three that are the same. Experimentally, this task is accomplished by training the animal to select the two objects that are the same (or, alternatively, the one that

is odd). The empirical question is whether the animal can learn to make the appropriate response over a diverse set of objects and then transfer that response to new instances. Such transfer implicates the learning of an abstract relation independent of the specific stimulus properties.

This task is like our explicit comparison task in that it asks the animal to find out what is the same about pairs of objects. And, it is a difficult one for many species, requiring many, many trials and even then transfer to new instances is not certain (e.g., Premack, 1978; Santiago & Wright, 1984). The very difficulty of this task supports our conclusion that explicit judgments of sameness require much more by way of internal representations than does mere categorization and response generalization. However, the fact that organisms without language *can* learn this matching-to-sample task suggests that language may not be *necessary* for the perceptual isolation and representation of abstract relations (see, especially, Oden, Thompson, & Premack, 1988).

Still, it may be that for children, language is the natural driving force behind the development of perceptual dimensions.

VIII. The Nontransparency of Mechanism

It is commonplace in psychological theorizing to impute to internal workings a "copy" of externally observed behavior (Smith & Thelen, 1993). Thus, sucking in infants is commonly explained by a sucking reflex, the alternate pattern of walking by a central pattern generator, categorization by represented categories, and syntax by an innate grammar. In each case, an abstract and sometimes truly elegant icon of the behavior to be explained is proposed to be the mechanism that produces the behavior. In cognitive psychology, this is particularly true when the behavior explained is verbal. Indeed, people's statements about their own cognitions have been argued to be the best windows on underlying category representation (Gelman & Coley, 1991; Keil, 1989). Moreover, one paradigm that has been used to study the relevant features for object categorization consists of *asking* people what the features are (e.g., Tversky, 1989; Rosch & Mervis, 1975; Malt, 1994; Rips, 1989). This, of course, is a far from foolproof method: The properties and features that people talk about could, in fact, have *no* localized internal representation as components. This point has also been empirically demonstrated by Schyns (1993). They showed that two groups of subjects could learn the very same three categories and perform equivalently in categorization tasks but have very different underlying representations. More specifically, the feature vocabularies defining contrasting categories varied qualitatively with the order in which the categories had been

learned. Learning the same three categories in different order created qualitatively distinct feature vocabularies that affected subsequent learning and judgments. In brief, categorizations by the subjects in this experiment were misleading indicators of the perceptual representations. The structure of outwardly observable behavior does not map in simple ways to the structure of the underlying mechanisms that make that behavior.

The nontransparency of mechanism is also demonstrated in the present work by the emergent and distributed nature of dimensional knowledge. Consider what the network knew about dimensions at the end of simulations 2 and 3. It knew, for example, what is the same about all red things. It knew that red and blue are attributes of the same kind, a kind different from that of big. It knew that *red* and *blue* are both possible answers to the question *What color is that?* and that *big* is not. All this knowledge was manifest in the collective behavior of the network. However, none of these bits of knowledge are located in any one place in the network. Rather, the system of dimensional knowledge displayed by the network is the emergent result of multiple mappings that overlap and constrain each other.

IX. Conclusion

The idea of perceptual dimensions as the primitive atoms of experience has figured prominently in the study of cognition (Berlin & Kay, 1969; Heider & Oliver, 1972; Lakoff, 1987; Miller & Johnson-Laird, 1976). The reason is clear. If perceptual dimensions are fixed and universal, then perception is a bedrock on which language, knowledge, and truth can be built. If, in contrast, what is perceived and therefore what is knowable from one's own interactions with the world and, from the language one learns, then there is no single truth. What is knowable is relative. This is a conclusion we may have to accept. The dimensions that structure our conscious experience of the world are themselves the product of experience.

REFERENCES

Aslin, R. N., & Smith, L. B. (1988). Perceptual development. In M. R. Rosenweig & L. W. Porter (Eds.), *Annual Review of Psychology, 39,* 435–474.

Backscheider, G. A., & Shatz, M. (1993, April). Children's acquisition of the lexical domain of color. Twenty-ninth regional meeting of the Chicago Linguistics Society, Chicago.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergency of symbols.* New York: Academic Press.

Berlin, B. D., & Kay, P. (1969). *Basic color terms, their universality and evolution.* Berkeley, CA: University of California Press.

Binet, A. (1980). Perceptions d'enfants. *Revue Philosphique, 30,* 582–611.

Boring, E. G. (1933). *The physical dimension of consciousness.* New York: Century.

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology.* New York: Irvington Publishers.

Bornstein, M. H. (1985). Colour-name versus shape-name learning in young children. *Journal of Child Language, 12,* 387–393.

Bruner, J. S. (1957). *Ongoing beyond the information given.* Cambridge, MA: Harvard University Press.

Callahan, M. A. (1990). Parent's description of objects: Potential data for children's inferences about category principles. *Cognitive Development, 5,* 101–122.

Carey, S. (1978). The child as a word learner. In M. Haller, J. Brenan, & G. Miller (Eds.), *Linguistic theory and psychological reality.* Cambridge, MA: MIT Press.

Carey, S. (1982). Semantic development. The state of the art. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art,* New York: Cambridge University Press.

Carnap, R. (1967). *The logical structure of the world and pseudoproblems in philosophy.* Berkeley, CA: University of California Press. (Original work published in 1928).

Clark, E. V. (1973). What's in a word: On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language.* New York: Academic Press.

Cohen, L. B., & Oakes, L. M. (1993). How infants perceive a simple causal event. *Developmental Psychology, 29,* 421–444.

Coldren, J. T., & Colombo, J. (1994). The nature and processes of proverbial learning: Implications from nine-month-old infants' discrimination problem solving. *Monographs of the Society for Research in Child Development, 59*(4), vii–75.

Cruse, D. A. (1977). A note on the learning of colour names. *Journal of Child Language. 4,* 305–311.

Donaldson, M., & Wales, R. J. (1970). On the acquisition of some relational terms. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 235–268). New York: Wiley.

Dromi, E. (1987). *Early lexical development.* New York: Cambridge University Press.

Ehri, L. (1976). Comprehension and production of adjectives and seriation. *Journal of Child Language, 3,* 369–384.

Fantz, R. L. (1963). Pattern vision in newborn infants. *Science, 140,* 296–297.

Freyd, J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics, 33,* 575–581.

Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

Gasser, M. S., & Smith, L. B. (in press). Learning nouns and adjectives: A connectionist account. *Language and Cognitive Processes.*

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). Cambridge, MA: Cambridge University Press.

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on thought and language: Interrelations in development.* Cambridge, MA: Cambridge University Press.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton-Mifflin.

Goldstone, R. L. (1995). Effects of categorization on color perception. *Psychological Science, 6*(5), 298–304.

Halff, H. M., Ortony, A., & Anderson, R. C. (1976). A context-sensitive representation of word meanings. *Memory & Cognition, 4*(4), 378–383.

Harnad, S., Hanson, S. J., & Lubin, J. (1991, March). Categorical perception and the evolution of unsupervised learning in neural nets. AAAI Spring symposium on Symbol Gounding: Problem and Practice, Stanford, CA.

Heider, E. R., & Olivier, D. C. (1972). The structure of color space in naming and memory for two languages. *Cognitive Psychology, 3,* 337–354.

Hurvich, L. M., & Jameson, D. (1957). An opponent-process theory of color vision. *Psychological Review, 64,* 384–404.

Husaim, J., & Cohen, L. (1981). Infant learning of ill-defined categories. *Merrill-Palmer Quarterly, 27,* 443–456.

Jackson-Maldonado, D., Thal, D., Marchman, V., & Bates, E. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language, 20,* 523–549.

James, W. (1890/1950). *The principles of psychology, Vol. 1.* New York: Dover Publications.

Karmiloff-Smith, A. (1986). From metaprocesses to conscious access: Evidence from children's metalinguistic and repair data. *Cognition, 23,* 95–147.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologists, 86,* 65–79.

Keil, F., & Carroll, J. (1980). The child's acquisition of "tall": Implications for an alternative view of semantic development. *Papers and reports on child language development, 19,* 21–28.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* Cambridge, MA: Cambridge University Press.

Kemler, G. G. (1982). The ability for dimensional analysis in preschool and retarded children: Evidence from comparison, conservation, and prediction tasks. *Journal of Experimental Child Psychology, 34,* 469–489.

Kendler, T. S. (1979). Cross-sectional research, longitudinal theory, and discriminative transfer ontogeny. *Human Development, 22,* 235–254.

Kotovsky, L., & Gentner, D. (1996). Progressive alignment: A mechanism for the development of relational similarity. *Child Development, 67,* 2797–2822.

Kruschke, J. K. (1992). ALCOVE: An exempler-based connectionist model for category learning. *Psychological Review, 99,* 22–44.

Locke, J. (1964). *An essay concerning human understanding.* New York: World Publishing.

MacIntosh, N. J. (1965). Selective attention in animal discrimination learning. *Psychological Bulletin, 64,* 125–150.

Malt, B. C. (1994). Water is not H_2O. *Cognitive Psychology, 27,* 41–70.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction.* Cambridge, MA: MIT Press.

Marks, L. E., Hammeal, R. J. & Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development, 51,* (1, Serial No. 215).

Marks, L. E., Szczesiul, R., & Ohlott, P. J. (1986). On the cross-modal perception of intensity. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 517–534.

Melara, R. D., & Mounts, J. R. W. (1994). Contextual influences on interactive processing: Effects of discriminability, quantity, and uncertainty. *Perception and Psychophysics, 56,* 73–90.

Mervis, C. B. (1987). Child-basic object categories and lexical development. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization.* Cambridge, England: Cambridge University Press.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception.* Cambridge, MA: Harvard University Press.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38,* (1-2, serial no. 149).

Oden, D. L., Thompson, R. K., & Premack, D. (1988). Spontaneous transfer of matching by infant chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 140–145.

Premack, D. (1976). *Intelligence in ape and man.* Hillsdale, NJ: Erlbaum.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning (Vol II).* New York: Appleton-Century-Crofts.

Rice, M. (1980). *Cognition to language categories: Word, meaning and training.* Baltimore: University Park Press.

Rips, L. J. (1989). Similarly, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21–59). Cambridge, England: Cambridge University Press.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Sandhofer, C. M., & Smith, L. B. (1996, November). *Learning a system of mappings: The acquisition of color terms.* Paper presented at the meeting of the Conference on Language Development, Boston, MA.

Sanocki, T. (1991). Intra- and interpattern relations in letter recognition. *Journal of Experimental Psychology: Human Perception and Performance.* 17(4), 924–941.

Sanocki, T. (1992). Effects of font- and letter-specific experience on the perceptual processing of letters. *American Journal of Psychology, 105*(3), 435–458.

Santiago, H. C., & Wright, A. A. (1984). Pigeon memory: *Same/different* concept learning, erial probe recognition acquisition, and probe delay effects recognition acquisition. *Journal of Experimental Psychology: Animal Behavior Processes, 10,* 498–512.

Schachter, P. (1986). Parts of speech systems. In T. Shopen (Ed.), *Language typology and syntactic description. Vol. 1. Clause structure.* Cambridge, MA: Cambridge University Press.

Schyns, P. (1993) Toward a theory of the development of primitives of representation in object categorization. Ph.D. Dissertation, Brown University.

Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (in press). *The development of features in object concepts. Brain and Behavioral Sciences*

Smith, L., & Sera, M. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology, 24,* 99–142.

Smith, L. B., & Kemler, D. G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology, 10,* 502–532.

Smith, L. B. (1984). Young children's understanding of attributes and dimensions: A comparison of conceptual and linguistic measures. *Child Development, 55,* 363–380.

Smith, L. B. (1993). The concept of same. *Advances in child development and behavior, 24,* 216–253.

Smith, L. B., & Evans, P. (1989). Similarly, identity, and dimensions: Perceptual classification in children and adults. In B. E. Shepp & S. Ballesteros (Eds.), *Object perception: Structure and process.* Hillsdale, NJ: Erlbaum.

Smith, L. B., & Thelen, E. (1993). Can dynamic systems theory be usefully applied in areas other than motor development? In L. B. Smith & E. Thelen (Eds.), *A dynamic systems approach to development: Applications* (pp. 151–170). Cambridge, MA: MIT Press.

Smith, L. B. (1989). From global similarities to kinds of similarities. The construction of dimensions in development. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogy,* (Pp. 146–178). Cambridge University Press

Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of Child Language, 4,* 1–22.

Stern, C., & Stern, W. (1924/1981). Die Kindersprache: einer psychologische und sprachtheoretische Untersuchung. Darmstadt: Wissenschaftliche Buchgesellschaft.

Thibaut, J-P., & Schyns, P. G. (1993). The development of feature spaces for similarity and categorization. *Psychologica Belgica, 35,* 167–185.

Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19,* 311–333.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 93–136.

Tversky, B. (1989). Parts, partonomies, and taxonomies. *Developmental Pscyhology, 25,* 983–995.

Werner, H. (1957). The concept of development from a comparative and organismic point of view. In D. B. Harris (Ed.), *The concept of development* (pp. 125–148). Minneapolis, MN: University of Minnesota Press.

Wohlwill, J. (1962). From perception to inference: A dimension of cognitive development. *Monographs of the Society for Research in Child Development, 72,* 87–107.

Wood, D. J. (1980). Teaching the young child. Some relationships between social interaction, language, and thought. In D. R. Olson (Ed.), *The Social foundations of language and thought.* New York: Norton.

Woodward, A. L., Markman, E., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology, 30,* 553–566.