COGNITIVE PSYCHOLOGY 24, 99–142 (1992)

A Developmental Analysis of the Polar Structure of Dimensions

LINDA B. SMITH

Indiana University

AND

MARIA D. SERA

Institute of Child Development
University of Minnesota

Some dimensions such as size and loudness are clearly marked for magnitude with one end being "more than" the other end. Thus, big is more than little and loud is more than quiet. This research examined the interaction of perception and language in the development of the magnitude marking of size, loudness, and achromatic color. Children 2 to 5 years of age and adults participated in six experiments. A cross-dimension matching task and a "Which is more?" task were used. The results suggest fast, invariant, and unidirectional development for size and loudness. Big is perceptually and linguistically more than little early in development and becomes more strongly organized with development. Loudness starts out disorganized and may be linguistically organized into more and less ends before it is perceptually organized. However, developments in perception and in language are rapid and in the same direction. In contrast, achromatic color shows an irregular developmental trend. Early in development, dark grey is perceptually more than light grey. But this early organization is disrupted at the same time that children acquire the words *dark* and *light*. The results suggest converging interactions between perception and language in the case of size and loudness and antagonistic interactions in the case of darkness. The results are interpreted in terms of a dynamical system.

A Developmental Analysis of the Polar Structure of Dimensions

Perceptual dimensions figure prominently in studies of language and thought (Berlin & Kay, 1969; Heider & Oliver, 1972; Lakoff, 1987; Miller & Johnson-Laird, 1976). The reason is clear. If perceptual dimensions are universal and uninfluenced by language, then perception is a bedrock on

This research was supported by NIH Grants R01-H719499 and K04-H1700589. We thank Jennifer Staab, Tracy Goodrich, Serina Johnson, and the Bloomington Developmental Learning Center. We thank Susan Jones, Dedre Gentner, Elissa Newport, and the reviewers for their comments on earlier versions of this paper. Reprint requests should be addressed to Linda B. Smith, Department of Psychology, Indiana University, Psychology Building, Bloomington, IN 47405.

0010-0285/92 $7.50

which language can be built. If, in contrast, what is perceived and therefore what is knowable from one's own interactions in the world depends on language, then there is no single truth. What is knowable is relative.

Many believe that the perception–language issue has been resolved in the favor of a constant and universal perceptual system (see Glucksberg, 1988). But the question of the psychological structure of perceptual dimensions and their relation to language is really a question about developmental process and here there is little relevant evidence. What is the role of perception and language in the *mechanism* of developmental change? What are the developmental points of dependencies and independencies between perceptual and lexical structure? Our goal in this work is the kind of data that would allow such questions to be answered. We seek an empirical description of the developmental *dynamics* of dimensional perception and dimensional language. We ask how the *speed, direction,* and *variability* of developmental change in perception and language are related. We specifically examined the development of the polar structure of quantitative dimensions.

More and Less Poles

Individual dimensions are logical systems and possess both syntax and semantics. The syntax of quantitative dimensions is the syntax of linear orders. A dimension is a linear order, if all its values fall on a line and if two directions can be defined with the following properties: (1) reciprocity—value A "is to the right of" value B if and only if B is "to the left of" A; and (2) transitivity—if A is "to the right of" B and B is "to the right of" C, then A is "to the right of" C. The semantics of dimensional structure concerns the psychological meaning of the directions. One direction of difference on a quantitative dimension is psychologically the direction of increase and the other direction is the direction of decrease. Thus, bigger is the direction of increase and littler is the direction of decrease in size. The semantics of quantitative dimensions is often conceptualized in terms of labeled poles. One pole is positive or "more" and movement toward that pole is an increase; the other pole is negative or "less" and movement in that direction is a decrease (see Clark, 1973; Holyoak, 1978).

The psychological independence of dimensional syntax and dimensional semantics is seen by comparing the dimensions of size, loudness, and achromatic color. All three dimensions are linear orders and exhibit the same logical properties. But the semantics of the three are not the same. Big is more than little. Loud is more than quiet. However, darker grey is not unambiguously more or less than lighter grey.

The common dimensional semantics of size and loudness and the contrasting semantics of achromatic color are seen in adults' performances in cross-dimension matching tasks. When asked to map sizes to loudnesses, all adults line up the two dimensions in one way: bigger is like louder and littler is like quieter. However, when adults are asked to map loudnesses to darknesses, half match louder sounds to lighter greys and half match louder sounds to darker greys (Marks, 1974, 1978). Thus, the semantics of size and loudness appear to be alike for all adults whereas the semantics of darkness varies.

The present research investigates the development of dimensional semantics. We ask what is it that the poles of size and loudness have that those of achromatic color do not? Where do the psychological meanings of a direction of increase and a direction of decrease come from?

The Problem of Origins

Many have suggested that the origin of the "more" and "less" ends of dimensions is to be found in the sensory system (see, for example, Boring, 1933; Clark, 1973; Marks, 1978; and Treisman & Gormican, 1988). Stevens (1957, 1975) argued for this view in his distinction between prothetic and metathetic dimensions. Stevens (1957) discovered that quantitative dimensions possess a unitary and well-ordered psychophysics and he called these dimensions "prothetic." In contrast, Stevens found that qualitative dimensions do not demonstrate a unitary or well-defined set of psychophysical properties and he called this disparate class of dimensions "metathetic." Stevens conjectured that the psychophysical properties that define the prothetic dimensions reflect a common physiology and that the myriad psychophysical properties of the metathetic dimensions reflect their distinct physiologies. In terms of their psychophysics, size and loudness are prothetic dimensions and achromatic color is a metathetic dimension (see Stevens, 1957, 1975). By Stevens's conjecture, then, size and loudness have a common sensory physiology that is not possessed by achromatic color. The implication is that directions of psychological increase and decrease are specified by the physiology of the sensory system.

Quantitative dimensions are distinguished by the linguistic structure of the words we use to talk about them as well as by their psychophysical properties (see Bierwisch, 1970; Bierwisch & Lang, 1989; Clark, 1973). For example, in English, the more-end term is also the neutral (or "unmarked") term by which we can refer to the dimension as a whole. *Big* and *loud* both possess the properties of more-end terms in English. None of the words used to talk about achromatic color—not *dark, light, black,* or *white*—clearly possess the properties of English more-end terms. In language as in psychophysics, then, size and loudness have unambiguous more-ends and less-ends and achromatic color does not.

Given these facts, a single seductive conclusion beckons: Language has the structure that it does because of perception and perception has the structure that it does because of the underlying sensory physiology. In our view, this conclusion is not warranted—at least not as a developmental account. The facts we have presented about size, loudness, and achromatic color are facts about the *mature* structures of the dimensions. Evidence about the endstate of mature perception and language is an insufficient basis for understanding how perception, cognition, and language interact in development to *make* that endstate. The endstate may be a poor guide to developmental origins and process.

In the following experiments, we ask where the "more" and "less" of quantitative dimensions come from by studying the development of two dimensions that possess "more" and "less" ends at maturity and one that does not. Specifically, we investigate the interaction of perception and language in the development of size, loudness, and achromatic color. We were encouraged in this enterprise by several anecdotal reports that very young children's perception of achromatic color is organized by psychological directions of increase and decrease (Marks, Hammeal, & Bornstein, 1987; Smith, 1987, 1989).

Developmental Dynamics and Between-Subject Variability

We seek to understand the origins of the more and less ends of dimensions by studying the developmental pathway that leads to mature organizations. Our approach differs from that typical of research in cognitive development and leads us to a different treatment of our data. The goal of most research in cognitive development is a description of the knowledge structures that are characteristic at different developmental levels. This focus on presumably stable and nonidiosyncratic cognitive structures promotes an empirical emphasis on the mean or best performance characteristic of a developmental level.

Our focus, in contrast, is on the dynamics of development. By dynamics, we mean the rate of change, the direction of change, and the variability of patterns of change. These are properties of the developmental path as a whole. Our goal is thus not a description of knowledge states at different developmental levels but a description of the shape of the developmental pathway itself. We attempt to infer the topology of the developmental pathway from cross-sectional data. We believe that just as developmental changes in mean performance have proven useful to those seeking theoretical descriptions of cognitive structures, *developmental changes* in *between-subject* variability provide important clues to the dynamics of development. In turn, we believe that the dynamic properties of development will provide new insights to interactions between perception and language.

The Experiments

The experiments examined the developmental dynamics of size, loudness, and achromatic color. To avoid misinterpretation, we distinguish here the dimensions of achromatic color and brightness. Brightness (illumination) and achromatic color (or surface darkness) are distinct perceptual dimensions. Brightness is a prothetic dimension by Stevens's psychophysical criteria whereas surface darkness is not (Stevens, 1957). Moreover, in cross-modal matching experiments, adults consistently map bright lights to loud sounds and dim lights to quiet sounds but they do not consistently map dark greys to loud sounds and light greys to quiet sounds (or vice versa, Marks, 1974; Marks, Szczesiul, & Ohlott, 1986). Finally, the phenomenon of darkness constancy shows that the perception of surfaces as white or black is not simply a matter of how much energy reaches the eye. Instead, dark objects remain constantly dark and light ones remain constantly light as the dimmer switch increases and decreases the overall illumination and the amount of reflected light. In this paper, we use the term "darkness" to refer to achromatic color and "brightness" to refer to illumination. The experiments examined only achromatic color (i.e., darkness) and size and loudness.

Across the six experiments, we employed two tasks. In Experiments 1, 2, 3, and 6, subjects were asked to map values on one dimension to values on another. If subjects represent two dimensions as each having a more end and a less end, they ought to perceive values on the more ends of the two dimensions as alike and values on the less ends of the two dimensions as alike. For example, if subjects perceive big to be more than little and loud to be more than quiet, than they should also perceive big to be like loud and little to be like quiet. In Experiments 4 and 5, we directly asked subjects to judge which end of a single dimension was the "more" end. In both kinds of tasks, cross-dimension matching and "more" judgments, we employed a perceptual version and a language version within subjects. Finally, our subjects were children from 2 to 5 years of age and adults. The age range in children comprises the period of most rapid growth in the perceived structure of dimensions, in dimensional concepts, and in dimensional language (for reviews, see Blewitt, 1982; Carey, 1982; Johnston, 1985; Smith, 1989).

EXPERIMENT 1

The specific questions of this experiment are whether children perceive big to be like loud (and little to be like quiet) and whether they perceive big to be like dark (and little to be like light or vice versa). The first case would be an instance of a perceived correspondence between two adult prothetic dimensions. The second case would be a perceived correspondence between an adult prothetic and an adult metathetic dimension.

Our specific cross-dimension matching task is structured as follows: The child is shown an exemplar that possesses an extreme value on one dimension. For example, the exemplar might be large. The child is then shown two choice objects that are both the same medium size but emit sounds that differ in loudness. The child's task is to indicate which of the two choice objects is most similar to the exemplar. If loud is perceptually like big, the child should pick the loud choice object as like the big exemplar. Four unique kinds of matches were required to document a complete magnitude alignment of size and loudness: (1) a match between a big exemplar and the louder of two sounds; (2) a match between a little exemplar and the quieter of two sounds; (3) a match between a loud exemplar and the bigger of two objects; and (4) a match between a quiet exemplar and the littler of two objects. Figure 1 shows the structure of these four trials for the combination of size–darkness and for the combination of size–loudness. Figure 1 also shows the structure of a parallel set of trials (the W-P trials) in which the exemplar was not presented but was described by a single word. These trials assess the degree to which dimensional *words* signal directions of increase and decrease that apply across dimensions.

Method

Subjects. A total of 24 children at each of three age levels participated: 2 year olds (2;3 to 2;11), 3 year olds (3;1 to 3;10), and 4 year olds (4;0 to 4;9). Half the children, equal males and females, participated in the Size–darkness condition and half participated in the Size–loudness condition. The children were recruited from birth announcements and advertisements in the local newspaper and were tested in the laboratory at the Psychology Department.

Stimuli and design. The stimuli were foam-core mice shaped as shown in Fig. 1. The half-circle portion of each mouse was covered with grey Coloraid paper. Tails were made from red pipe cleaners and the eyes and whiskers were pink. The mice were capable of standing upright and were always presented in that manner.

The diameters of the half-circle were 4.5, 7.5, and 11.5 cm. In the Size–darkness condition, the grey colors of the mice were a light grey—Coloraid No. 2, a midgrey—Coloraid No. 5, or a very dark grey—Coloraid No. 8. In the Size–loudness condition all mice were midgrey. The sound in the Size-loudness condition consisted of two 700-ms repetitions of a 610-Hz tone separated by 250 ms of silence. This sound was recorded on a Marantz tape recorder. The quiet sound was played at 53dB SPL and the loud sound was played at 82dB SPL. No sounds were associated with the mice in the Size–darkness condition. Two 30-cm^2 houses made of brown cardboard and decorated with red doors and windows were also employed. The following adjectives were also used: *big, little, loud, quiet, dark, light.*

Tasks and design. There were three tasks (see Fig. 1): (1) *The Percept-to-Percept* (*P-P*) *task* asks how a polar value of one dimension maps onto the polar values of another dimension; the child is presented with an exemplar stimulus of maximal or minimal value on one dimension and is asked which of two choice stimuli of maximal and minimal values on the *other* dimension is like the exemplar. (2) The *Word-to-Percept* (*W-P*) *task* asks how the *words* about one dimension map onto the perceived poles of the other dimension; the child is presented with a polar adjective that describes a *not* present mouse and is asked which of

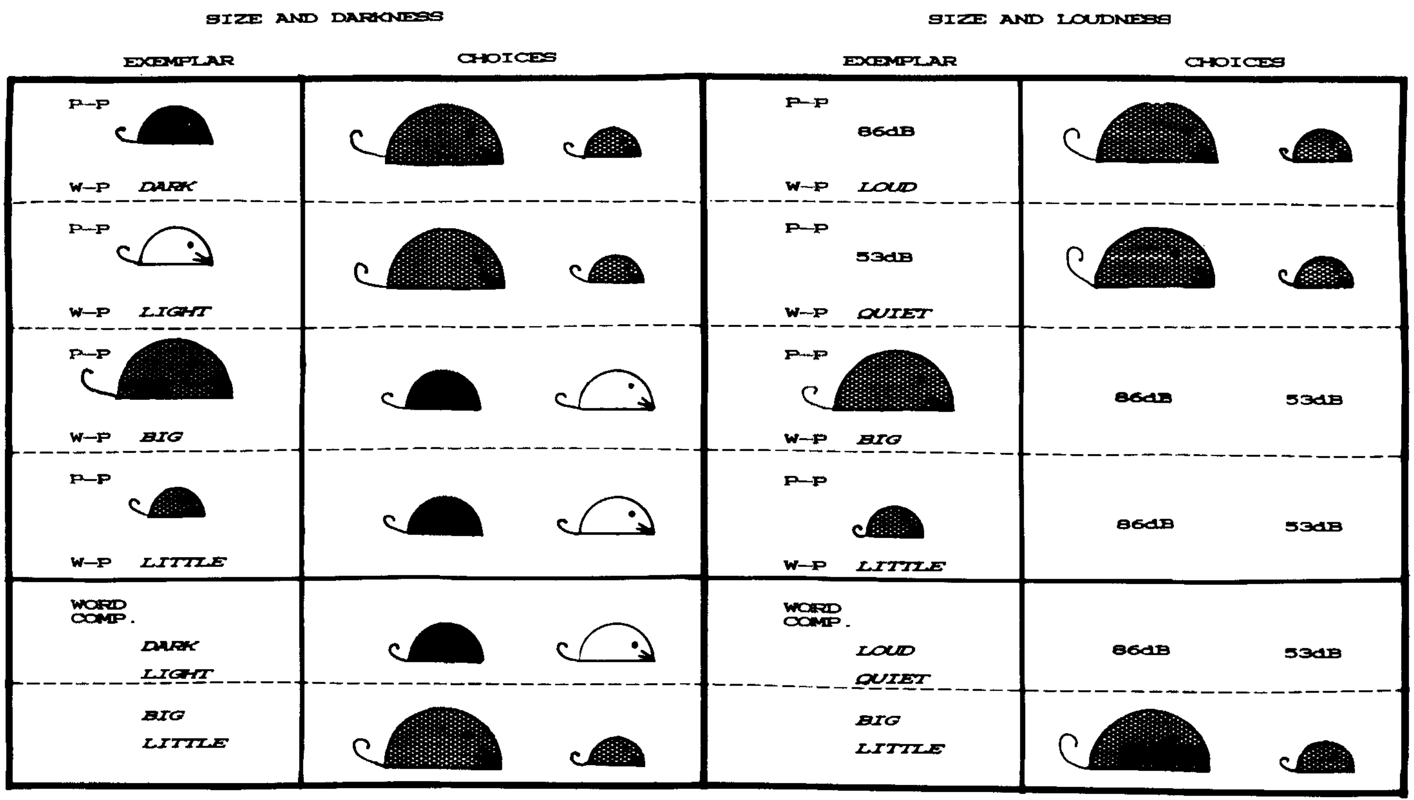

FIG. 1. Illustrations of the size–darkness trials and the size–loudness trials of Experiment 1. P-P labels the exemplars for the Percept-to-Percept task. W-P labels the exemplars for the Word-to-Percept task. The bottom two rows show the trials for the Word Comprehension task.

the perceptually present choice stimuli that vary on the nondescribed dimension is like the described mouse. (3) The *Word-comprehension task* assesses word knowledge; the child is asked which of two objects is described by an adjective.

The unique trials that made up each condition of dimensional combination are depicted in Fig. 1. The four unique P-P and W-P trials for each condition differ from each other only in the exemplar (object or word). Each of these trials was repeated four times for a total of 16 P-P and 16 W-P Size–darkness trials and 16 P-P and 16 W-P Size–loudness trials. The four unique Word-comprehension trials were repeated twice for a total of 8 trials in each condition. The trials within each task were presented in a randomly determined order.

The Size–darkness and Size–loudness conditions varied between subjects. Task varied within subject. The order of the P-P and W-P tasks were counterbalanced and the two tasks were conducted in separate sessions within a 2-week period. The Word-comprehension task was presented in the session with the W-P task and order of tasks within that session was counterbalanced across subjects.

Procedure. The child was first shown the entire range of mice and in the Size–loudness condition was exposed to the two sounds that would be used. The mice were removed before test trials began.

The *Size–darkness P-P trials* were structured as follows: The exemplar mouse was placed in front of a house. The child was told that the mouse lived in the house and that a friend who *was very much like the exemplar mouse*—who played the same games, liked the same things—lived in the second, next door, house. The two Choice mice were then introduced. The exemplar mouse remained in view and the child was asked to select the friend "who was very much like" the exemplar. All subsequent trials were structured in the same way.

In the *Size–darkness W-P task,* the exemplar mouse was verbally described but was not visually present. The experimenter began a trial by motioning to one house and saying "a mouse lives in this house and he is *big* (*little/dark/light*) and he has a friend who is very much like him—who likes to play the same games and likes the same things." The child was then asked to pick which of the two visually present choices was like the "big mouse."

In the *Size–darkness Word comprehension task,* the child was presented with two mice that varied either in size or darkness and asked "Which is the *big/little/dark/light* one." The term queried was appropriate to the dimension of variation.

Since two stimuli with different sounds cannot be presented simultaneously, the details of the size–loudness tasks do not correspond perfectly with their counterparts in the size–darkness condition.

In the *Size–loudness P-P task,* the two cardboard houses were separated by 30 cm. The trials in which the exemplar was a mouse and the choice stimuli were sounds and the trials in which the exemplar was a sound and the choice stimuli were mice were structured somewhat differently. On the trials in which the exemplar was a mouse, the exemplar mouse was placed equidistant from the two houses. The child was told that this mouse was looking for his friend "who was very much like him (etc.)." The child was told that the friend was in one of the two houses. The child was told to listen to the sounds coming from the houses to figure out which was the friend's house. A tape recorder located behind one house then played the tone at its designated loudness. Then a tape recorder located behind the second house played its tone at its designated loudness. The child was reminded that the exemplar mouse was looking for his friend who was very much like him. The child was asked to listen again as the two tones were replayed. The child then indicated which sound came from the mouse most like the exemplar mouse. The sounds were repeated if the child hesitated or if the child asked to hear them again. The order of playing the loud and quiet sounds was counterbalanced across trials.

On the Size–loudness P-P trials in which the exemplar was a sound, the child was told that the exemplar mouse was in his house calling for his friend. The appropriate sound from a recorder behind one house was played and the child chose from the visually presented choices. The left–right position of the loud and quiet sounds varied across children.

The *Size–loudness W-P and Word comprehension tasks* were analogous to those in the Size–darkness condition. In the size–loudness W-P task when the not present "exemplar" was described by a size term, the child chose between two sounds that emanated from the houses; no choice mice were in view. When the not present "exemplar" was described by a loudness term the child chose from two mice differing in size. In the word-comprehension task, when size terms were tested, the child chose from two mice differing in size. When loudness terms were tested, the child chose from sounds with no mice in view. On these trials, each sound was played, the two sounds emanated from distinct locations marked by the houses. The child was then asked "Which is loud?" (or on half of these trials, "Which is quiet?") "Listen." The two sounds were played again and the child was asked to respond by pointing to the house. The sounds were repeated as often as necessary.

Results and Discussion

We report separately the results for the Size–loudness and Size–darkness conditions beginning with the Size–loudness data because they are the most straightforward.

Size and Loudness

Children's mean polar-consistent choices are shown in Fig. 2. We defined polar-consistent choices as big–loud/little–quiet alignments of the two dimensions. Children's polar-consistent choices were submitted to an analysis of variance for a 3 (Age) × 2 (Task) × 2 (Pole) × 2 (Dimension serving as the exemplar) mixed design. The factor labeled "pole" refers to whether the exemplar object or adjective was from the "more" or "less" end of the dimension. The analysis revealed a reliable main effect of Age, $F(2,33) = 62.99$, $p < .001$. As is obvious from Fig. 2, 3- and 4-year-old children made many more choices that matched the more and less ends of the two dimensions than did the 2 year olds. The analysis also revealed a main effect of Task, $F(1,33) = 7.17$, $p < .02$, and a reliable Age

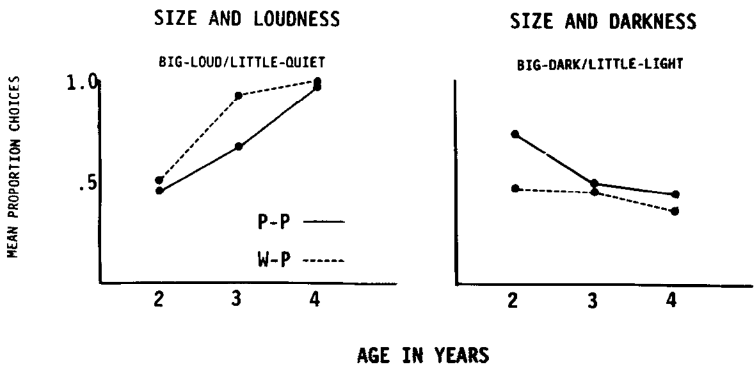

FIG. 2. Mean proportion pole-consistent mappings in the P-P and W-P tasks.

× Task interaction, $F(2,33) = 5.73, p < .02$. Post hoc analyses (Tukey's h.s.d., $\alpha < .05$) indicated big–loud/little–quiet matches were more common in the W-P task than in the P-P task for the 3 year olds. No other main effects or interactions approached significance.

The left side of Table 1 shows the mean performances of children in the Word-comprehension task. The number of children who performed perfectly in this task for each word (i.e., chose correctly two out of two times) are given in parentheses. Almost all children, including the 2 year olds, understood the words *big* and *little*. The 3 and 4 year olds also understood the words *loud* and *quiet* but 2 year olds did not. The level of understanding of the sound terms is less than is suggested by the numbers in Table 1. Four of the 2 year olds chose the louder sound both when asked to select the *loud* sound and when asked to select the *quiet* one; three 2 year olds used the opposing strategy—always selecting the quieter sound. Only two of the twelve 2 year olds chose correctly on both the *loud* and *quiet* Word-comprehension trials.

Size and Darkness

We scored children's size–darkness judgments in terms of a big–dark and little–light alignment. The mean choices consistent with this cross-dimension mapping are also shown in Fig. 2. The opposing correspondence (i.e., a mapping of big to light and of little to dark) is given by 1 minus the mean proportion in the figure. We submitted children's number of choices consistent with the designated mapping to an analysis of variance for a 3 (Age) × 2 (Task) × 2 (Pole—with dark designated as the positive pole) × 2 (Dimension serving as the exemplar) mixed design. The analysis revealed only reliable main effects of Age, $F(2,33) = 3.47, p < .05$, and Task, $F(1,33) = 5.72, p < .05$. As is evident in the figure and as confirmed by post hoc tests (Tukey's h.s.d., $\alpha < .05$, h.s.d.), 2 year olds

TABLE 1
Mean Proportion Correct Responses in the Word–Comprehension Task of Experiment 1

	Size–loudness				Size–darkness			
	Big	Little	Loud	Quiet	Big	Little	Dark	Light
2 years	.85	.87	.62	.68	.83	.83	.75	.52
	(10)	(10)	(06)	(05)	(09)	(10)	(07)	(04)
3 years	1.00	.96	.96	.96	.96	.94	.88	.60
	(12)	(11)	(11)	(11)	(11)	(10)	(09)	(07)
4 years	.96	1.00	1.00	1.00	1.00	.96	.88	.77
	(11)	(12)	(12)	(12)	(12)	(10)	(10)	(09)

Note. Number of children (max = 12) choosing correctly on 2 out of 2 trials is given in parentheses.

made more big–dark/little–light choices than did older children. The Age × Task interaction did not approach significance because of the extreme variability among the performances of older children, a point we discuss below. However, the consistent maps by 2 year olds appear restricted to the P-P task. Thus, if we ask of the group performances whether children's choices in a task deviated from chance, the answer is yes only for 2 year olds in the P-P condition, $t(11) = 5.08, p < .01$. For 2 year olds as a group in the percept-to-percept task only, big and dark are alike and little and light are alike. This result confirms previous anecdotal reports (Marks et al., 1987; Smith, 1989).

The right side of Table 1 shows the mean performances in the Word-comprehension task. The numbers of children showing perfect comprehension of each term are given in parentheses. As did their counterparts in the Size–loudness condition, most children in the Size–darkness condition, including the 2 year olds, correctly understood the terms *big* and *little*. However, children's comprehension of the words *dark* and *light* as applied to achromatic color appears to grow much more slowly during this period. The numbers of children understanding both terms correctly (i.e., scoring four for four correct across both color terms) are one, eight, and seven for 2, 3, and 4 year olds, respectively.

State Spaces

How are changes in the perceptual structure of dimensions related to changes in understanding dimension words? We use scatterplots of individual performances in state spaces defined by the P-P and W-P task to examine developmental changes in between-subject variability simultaneously across the two tasks. These scatterplots are shown in Fig. 3.

We explain these state spaces by considering the graph for size and loudness which is shown on the left. On the y axis is the proportion of big–loud/little–quiet matches in the W-P task. On the x axis is the proportion of big–loud/little–quiet choices in the P-P task. The individual symbols indicate the performances of individual children. Thus, a child who matched the word *big* to loud, the word *little* to quiet, the word *loud* to big, and the word *quiet* to little with perfect consistency in the W-P task *and* matched big to loud, little to quiet, loud to big, and quiet to little in the P-P task would fall in the upper right-most portion of the state space. A child who responds randomly in a task would be expected to make big–loud/little–quiet matches on .50 of the trials of that task. Thus, a child who responds randomly in both tasks would fall in the center of the state space.

The proportion of choices consistent with an "opposite" alignment of the two dimensions (i.e., big–quiet and little–loud) is given by 1 minus the proportion in the figure. A child who always matched big to quiet,

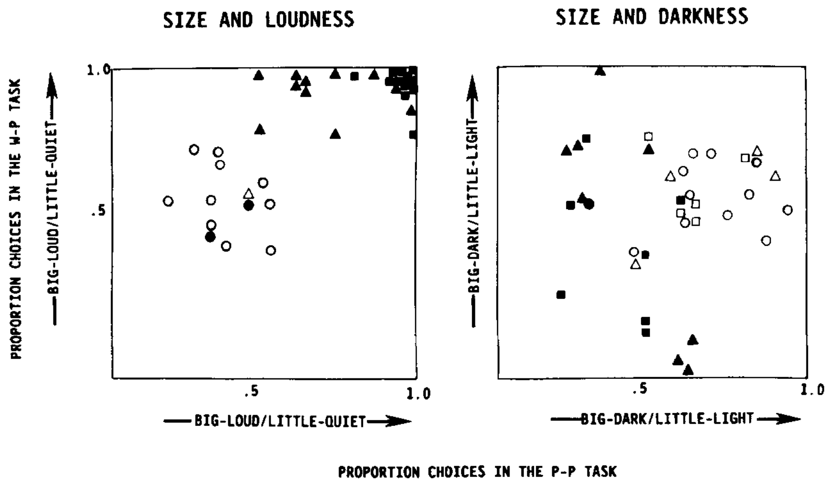

FIG. 3. Each individual's designated choices in the W-P task are plotted against their choices in the P-P task. Individual 2 year olds are indicated by circles, 3 year olds by triangles, and 5 year olds by squares. Solid symbols indicate perfect comprehension of the dimensional terms.

quiet to big, little to loud, and loud to little in *both* the W-P and P-P tasks would be indicated by a symbol in the lowest left-hand portion of the state space. As is apparent, no child was this perverse. In sum, individual children's performances are indicated by their location in the state space. Children who make big–loud/little–quiet choices in both the W-P and P-P tasks are located in the upper right quadrant of the space. Those who make the opposite alignment of big to quiet and little to loud in both tasks would be located in the lower right quadrant. Deviations from chance (from the center) in the direction of the remaining two quadrants indicate discordant alignments of the dimensions in the percept and word tasks. Solid symbols designate children who showed perfect comprehension of all four adjectives (i.e., eight out of eight correct responses) in the Word-comprehension task.

This representation of the data shows clear growth in the relation between performances in the P-P, W-P, and Word-comprehension tasks in the Size–loudness condition. The 2 year olds' performances, indicated by circles, are scattered about the center of the state space. The center is the region corresponding to chance level performance in both the P-P and W-P tasks. The circles are mostly unfilled, indicating that children made errors on at least some terms (most typically the loudness terms) in the Word-comprehension task. The 3 and 4 year olds are designated by triangles and squares. These symbols fall in the upper right of the state space—in the region which corresponds to big–loud/little–quiet mappings in both the W-P and P-P tasks. The symbols indicating individual 3 and 4 year olds are also mostly filled as most older children perfectly understood the words *big, little, loud,* and *quiet* in the Word-comprehension task.

The right side of Fig. 3 shows the scatterplot of individual performances in the state space defined by the P-P and W-P tasks in the Size–darkness condition. Each individual's big–dark/little–light matches in the W-P task are plotted as a function of that individual's choices in the P-P task. Again, chance in each task is .50 and the proportion of choices consistent with an opposing mapping of the two dimensions (i.e., big–light and little–dark) is given by 1 minus the proportion in the figure. Thus, deviations from chance (the center) toward the upper right quadrant indicate big–dark/little–light choices in both tasks and deviations from center toward the lower left quadrant indicate big–light/little–dark choices in both the P-P and W-P tasks. Individuals in the upper left and lower right of the state space had discordant patterns of cross-dimension matches in the P-P and W-P tasks. For example, children in the lower right of the space are children who matched the word *big* to light, the word *little* to dark, the word *dark* to little, and the word *light* to big but matched big mice to dark mice (and vice versa) and little mice to light mice (and vice versa). Again, subjects indicated by solid symbols are children who showed perfect comprehension of the four adjectives (i.e., eight out of eight correct responses) in the Word-comprehension task.

The performances of 2 year olds and other children who do not fully understand all the dimensional terms (all open symbols) are localized in one region of the state space. This region corresponds to high proportions of matches of big with dark and little with light in the P-P task but chance level performance in the W-P task. Chance level performance on the W-P task is, of course, not surprising since many of these children do not understand the words as indicated by the open symbols.

Older children, many of whom also understand *all* the dimensional terms (the solid triangles and squares), differ widely as to how they align size and darkness in the two tasks. Some individual children exhibit consistent alignments in one task that are at odds with their alignments in the other task. For example, several 3 year olds maintained a mapping that linked big with dark and little with light on 60% of the P-P trials but maintained the inverse map of linking big with light and little with dark on over 75% of the W-P trials. Other children were consistent in mapping sizes to darknesses in one direction or the other in one task but not the other. Still other children were inconsistent in their mappings of sizes to darknesses in both tasks. The best characterization of the older children's performances is that there is no single pattern—across older children or across the perception and language tasks.

In sum, very young children with weak knowledge of the words for contrasting darkness exhibit a single pattern of performance—big is perceptually like dark and little is perceptually like light. With development and at about the same time that the words *dark* and *light* are understood, there is a disruption of this early perceptual correspondence and a dispersion of individuals throughout the state space.

Developmental Trajectories

We estimated individual developmental trajectories from the present cross-sectional data by randomly connecting (without replacement) each 2-year-old data point to a 3-year-old data point and then randomly connecting (without replacement) each of these 1-year trajectories to a 4-year-old data point. These hypothetical developmental pathways are shown in Fig. 4. Our procedure of random connection potentially exaggerates the disorderliness and variability of real individual trajectories because these trajectories connect different individuals.

Nonetheless, the suggested developmental pathway for the mapping between size and loudness appears relatively narrow and orderly. Development starts in the center of the state space: individuals do not consistently align sizes and loudnesses in any particular way. But there is rapid developmental movement in one direction—to a consistent mapping of the more and less ends of the two dimensions and to a corresponding cross-dimension map between dimension words and perceived values on a dimension. The developmental pathway may bend upward a bit in the state space—advancing first in the task involving the mapping of words to

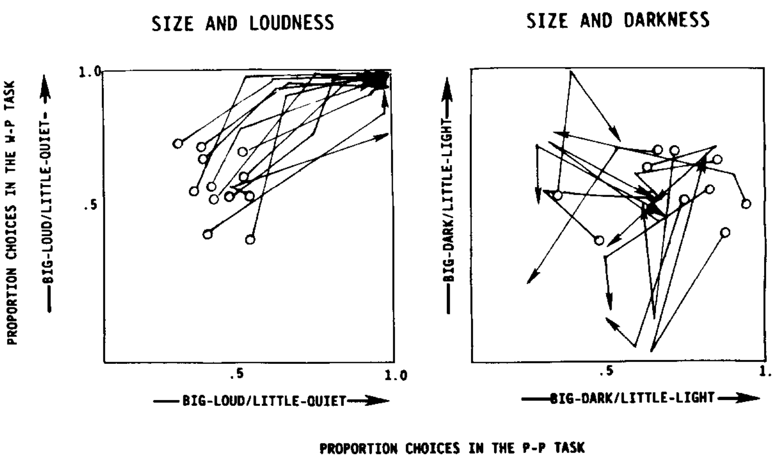

FIG. 4. Estimated developmental trajectories for the data shown in Fig. 3.

percepts and subsequently in the task involving the mapping of percepts to percepts. Overall, the data suggest rapid, direct development to one endpoint.

The estimated trajectories for size and darkness are far more disorderly and do not suggest a well-defined or singular pathway. Development proceeds from a localized starting point that maps perceived sizes onto perceived darknesses. But from this starting point in perception there is with development a scattering of individuals. Our procedure of randomly connecting individuals at different development levels may exaggerate the disorderliness of the developmental pathway and may suggest individual trajectories that do not exist. However, given the similarity of individual 2 year olds to each other and the dissimilarity of individual 4 and 5 year olds to each other, all possible connection schemes result in a picture of development in which there is chaotic movement in several directions *away* from an initial starting point.

Our next experiment was designed to replicate these findings and to address two questions. First, is the greater difficulty of the Size–loudness condition than the Size–darkness condition for 2 year olds related to the differences in task procedures in the two conditions? In the Size–loudness condition of Experiment 1, children matched disembodied sounds to objects. Young children, in particular, may not have understood their task. Although it is not possible to fully equate a cross-modal matching task to a within-modality matching task, Experiment 2 attempts to minimize the differences between the Size–loudness and Size–darkness conditions. Second, in order to obtain a fuller picture of the developmental pathway, we included 5 year olds and adults in Experiment 2.

EXPERIMENT 2

Method

Subjects. A total of 24 subjects at each of five age levels participated: 2 year olds (2;3 to 2;10), 3 year olds (3;1 to 3;9), 4 year olds (4;2 to 4;11), 5 year olds (5;1 to 5;11), and college undergraduates. Half the subjects at each age level, equal males and females, were randomly assigned to the Size–darkness condition, and half were assigned to the Size–loudness condition.

Stimuli and procedure. The stimuli were identical to those used in Experiment 1 with the exception that a third sound of intermediate loudness (68 dB) was added and all mice sat on a small box that contained a tone generator set at its assigned loudness and calibrated daily. The sound (of set duration and pitch, see Experiment 1) was turned on by a button switch on the box.

The procedure and design were identical to those of Experiment 1 except that in both the Size–loudness and Size–darkness conditions, all P-P trials consisted of three mice, all of whom made a noise. This was done to more fully equate the stimuli in the Size–loudness and Size–darkness conditions. In the Size–darkness condition, all mice made the sound of intermediate loudness just as in the Size–loudness condition all mice were the intermediate value of darkness. All trials in the W-P task in both conditions consisted of an adjective and two choice mice, both of whom emitted sounds. Again, the choice mice in the Size–darkness condition all emitted the sound of middle loudness just as the choice mice in the Size–loudness condition were midgrey.

The procedure in a session was the same for children and adults. The adults were told before the experiment began that this task was being used with 2 year olds and that, for comparison purposes, they were being tested with the same instructions as 2 year olds.

At the start of each session, the subject was shown the full range of mice and was shown how to make each one "talk." On each P-P trial, the subject was introduced to the exemplar mouse and the experimenter pushed the button causing the sound to be played. The subject was told that the exemplar mouse was looking for his friend who was very much like him (etc.). The two choice mice were placed on the table. Each of their sounds was played. Then the subject was asked "Which one is his (her) friend? Listen, they are talking." Each sound was then replayed in this order: the exemplar sound, the first choice mouse, and the second choice mouse. (Again, the sounds varied only in the Size–loudness condition.) The subject was then asked "Which mouse (motioning to the choices) is most like this mouse?" (motioning to the exemplar). The procedure in the W-P task was identical except that the exemplar mouse was not in view and made no sounds but was instead described by a single adjective. All other aspects of the experiment were identical to that of Experiment 1 with the additional exception that there was no Word-comprehension task.

Results and Discussion

Again, we begin with the results from the Size–loudness condition.

Size and Loudness

Subjects' choices consistent with a big–loud/little–quiet mapping were submitted to an analysis of variance for a 5 (Age) × 2 (Task) × 2 (Pole) × 2 (Exemplar dimension) mixed design. The analysis revealed only reliable main effects of Age, $F\ (4,55) = 24.84$, $p < .001$, and Task, $F\ (1,55) = 4.042$, $p < .05$. Figure 5 shows the mean proportion of choices. Big–loud/little–quiet matches increased dramatically in children from 2 to 4 years of age and such matches were more prevalent for children of all ages in the W-P than in the P-P task.

Figure 6 (top) shows scatterplots of individuals' cross-dimension matches in the W-P and P-P tasks separately for each age level. As in Experiment 1, 2 year olds' performances are scattered about the center of the state space—the region corresponding to chance level mappings of sizes and loudnesses in both tasks. Three year olds' performances are also scattered but many of them are located in the region of consistent big–loud/little–quiet mappings. The fact that 3 year olds' performances were more variable in this experiment than in Experiment 1 suggest that contrary to our goal we may have increased the perceptual complexity of the task. However, in both the P-P and W-P tasks, older children and adults uniformly judged big to be like loud and little to be like quiet. Despite the minor differences in results of this experiment and Experiment 2, the emerging picture of development is the same: rapid progres-

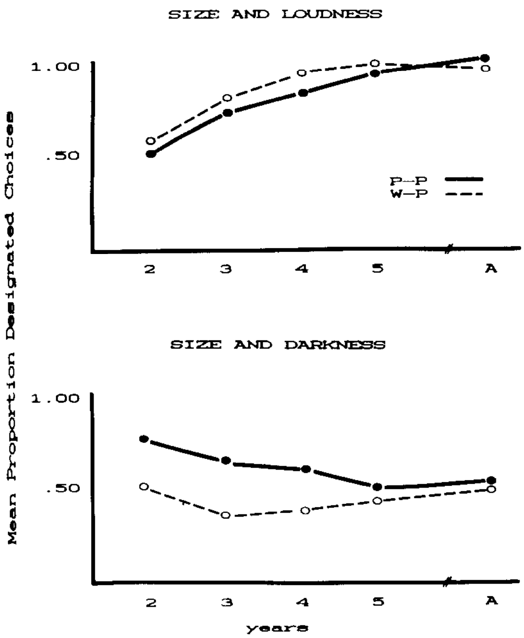

FIG. 5. Mean proportions of big–loud/little–quiet mappings and mean proportions of big–dark/little–light mappings in the P-P and W-P task at the five age levels.

sion to a perceptual and linguistic correspondence between big and loud and between little and quiet.

Size and Darkness

Subjects' choices consistent with a big–dark/little–light mapping were submitted to an analysis of variance for a 5 (Age) × 2 (Task) × 2 (Pole with dark and big designated as the positive poles) × 2 (Exemplar dimensions) mixed design. The analysis revealed only a main effect of Task, $F(1,55) = 22.95, p < .001$, and a reliable interaction between Age and Task, $F(4,55) = 3.37, p < .02$. Figure 5 (bottom) shows the mean proportion of big–dark/little–light choices. (The proportion of maps in the opposite direction, big–light/little–dark, are given by 1 minus the proportion in the figure.) As in Experiment 1, big–dark/little–light mappings occurred only in the P-P task and were made only by the youngest subjects. 2 year olds aligned size and darkness in one way with a consistency that is reliably above chance, $t(11) = 6.96, p < .001$. At no other age level in the P-P task and at no age level in the W-P task did group performances deviate from chance. This is not surprising given the large increase with age in between-subject variability in both tasks.

Figure 6 (bottom) shows the scatterplots of individual performances at

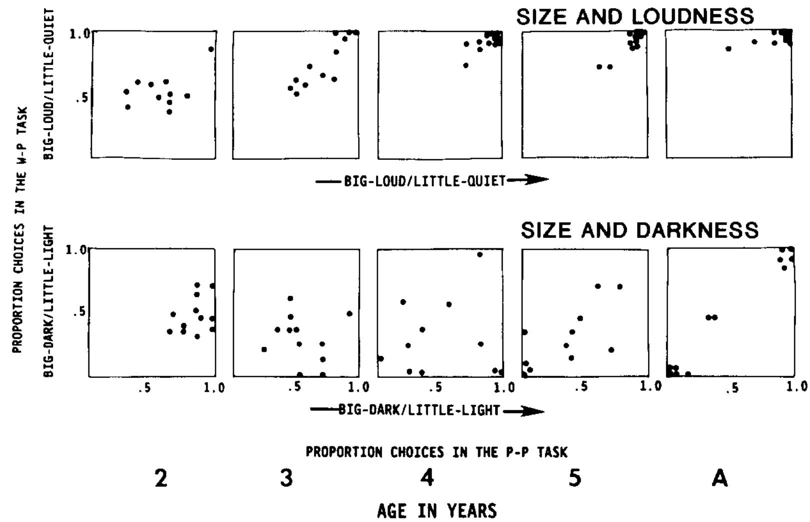

FIG. 6. Each dot indicates the performance of one subject. In the top figures each individual's proportions of big–loud/little–quiet choices in the W-P task are plotted as a function of big–loud/little–quiet choices in the P-P task for each age level separately. In the bottom figures each individual's proportion of big–dark/little–light choices in the W-P task is plotted as a function of big–dark/little–light choices in the P-P task.

each age level in the P-P and W-P task. Most 2 year olds, as did those in Experiment 1, consistently map big to dark and little to light in the P-P task. This cross-dimension similarity, however, does not carry over to 2 year olds' word-to-percept judgments which vary about the center of the state space, that is, about chance.

The performances of 3 and 4 year olds are highly variable in both tasks. Individual 4 year olds can be found to fit all possible patterns of correspondences. There is one who mapped big to dark and little to light with high consistency in both the P-P and W-P tasks. There are some who maintained the opposite mapping in both tasks. There are several who maintained a big–dark/little–light map in the percept task but maintained the *opposite* big–light/little–dark map in the word task. Children who made discordant maps in the W-P and P-P tasks are indicated by the dots in the upper left and lower right quadrants of the state spaces. In all, ten 3 and 4 year olds made discordant cross-dimension matches in the P-P and W-P tasks. At these ages, the cross-dimension similarities between percepts and between words and percepts are not necessarily the same.

For 5 year olds and more markedly for adults, performance is localized in one of three regions in the state space: the region in which big goes with dark and little goes with light in both the W-P and P-P tasks (upper right quadrant), the region in which big goes with light and little goes with dark in both tasks (lower left quadrant), and the region in which the two dimensions are not consistently related in both tasks (center).

The developmental trend for size and darkness thus appears to proceed from an organized state of a *perceptual* correspondence to disorganization and then to three possible states in which the correspondence between size and darkness varies across individuals but is organized within an individual in the same way in language and perception.

Developmental Trajectories

The principal results of the first two experiments are summarized by the trajectories in Fig. 7. The state space is defined as in Fig. 4. These trajectories, in contrast to those in Fig. 4, are suggested by the data but not directly estimated from the data. These two emerging pictures of the two developmental pathways suggest two distinct kinds of perception–language interactions. Size and loudness show orderly and rapid development to one unified organization of cross-dimension similarities. Language does not lag behind perceptual development, but may, instead, be a facilitating factor in cross-modal matches between size and loudness. In contrast, size and darkness show an early perceptual organization, disorganization, and then reorganization into three patterns of individual differences. Language seems initially to play a destabilizing role. When children first acquire words, words and percepts can lead to opposing cross-dimension correspondences. But later in development, amidst wide individual differences in the alignment of sizes and darknesses, a tight organization of perception and language emerges within the individual —at least in two highly similar tasks presented several days apart.

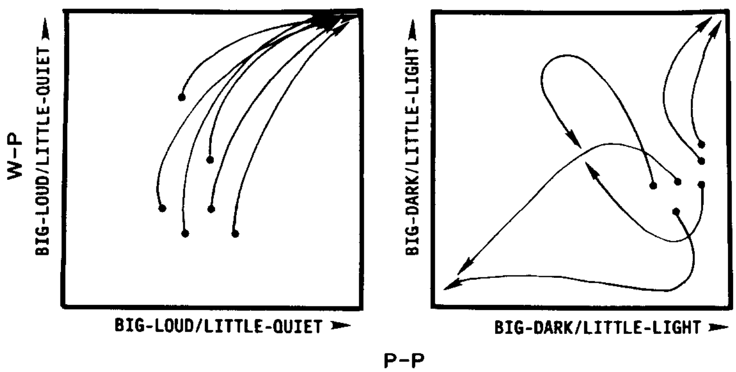

FIG. 7. Theoretical trajectories of developmental paths in the state space relating big–loud/little–quiet mappings in the W-P task to the state space relating to those in the P-P task on the left and big–dark/little–light mappings in the W-P task to those in the P-P task on the right.

EXPERIMENT 3

This experiment has two purposes. First, it replicates the developmental trends observed in Experiments 1 and 2 in a new cross-dimension matching task, and second, it addresses a specific question about cross-dimension correspondences. Specifically, are an individual's cross-dimension correspondences logically coherent? Are they transitive? If an individual judges big to be like loud *and* loud to be like light, does that individual then judge big to be like light?

In the experiment, children and adults were asked to align the poles of three combinations of dimensions—size–loudness, size–darkness, and loudness–darkness. Each individual made all three kinds of judgments. In contrast to the first two experiments, this experiment examined only the perceived similarities across the dimensions; there was no corresponding language task.

Method

Subjects. Sixteen individuals, equal males and females, participated at each of three age levels: 2 and young 3 year olds (2;6 to 3;3), 4 and 5 year olds (4;0 to 5;5), and adults. All subjects were tested in the laboratory.

Stimuli. The stimuli were cubes varying in size and darkness and the same auditory stimuli used in Experiment 1. The three possible side lengths of the cubes were 5, 7.5, and 10 cm. Each cube was painted one of three possible shades of grey: light, medium, and dark. The tones were recorded and played on a Marantz tape recorder at three loudness levels: 53, 68, and 82 dB SPL. For the children, the cubes and sounds were presented as pairs. Stimuli within a pair varied extremely on one dimension. For adults, cubes and sounds were organized into triplets of items that varied on one dimension.

Procedure. Slightly different procedures were used for the children and adults. The children's procedure began with three training trials. On the first training trial, the child was given a plate and a kettle. A toy fried egg and the cover to the kettle were placed on the table and the child was asked to "put these where they belong." If the child did not immediately put the egg on the plate and the cover on the kettle, the experimenter showed the child the appropriate responses. On the second training trial, the child was given a horse and a pail and then shown a saddle and a shovel and told to "put these where they belong." All children did so correctly. On the third training trial, the child was shown a toy cow and a baby doll. The experimenter then played on the tape recorder a "moo" and a baby cry. She then replayed the "moo" and asked "who does this sound go with?" All children indicated the appropriate objects.

Children were then given the first test trial. The procedure for these test trials is clarified by considering the structure of a trial in which the child is asked to match cubes varying in size to cubes varying in darkness. This trial began with the presentation of a dark and a light cube. The experimenter then presented one object from the pair to be matched—e.g., the big cube. The child was asked "where does this (the big cube) go? Does it go with this one (e.g., the dark one) or does it go with this one (e.g., the light one)?" After the child indicated a match, the remaining object, the little cube,was brought into view. While the big object remained spatially close to the indicated choice, the same questions were asked of the little cube. Children were not allowed to match two stimuli to a single object. If they tried to, for example, if they placed the big cube and the little cube next to the dark one, they were asked "Which one goes best with this one? Does this one (the one not chosen as best) go here then (motioning to the object without a match)?" Very few children ever attempted to match two objects to a single standard and none did so more than once.

All pairings of dimensions were presented four times. Which dimension in a pair was displayed first was counterbalanced across the size–darkness trials. For the size–loudness and darkness–loudness trials, the cubes (i.e., the visual displays) were always presented first and children were asked to match the sounds to the cubes. These trials were structured identically to the "cow-and-baby" training trial described above. Which pole of a contrasting pair of stimuli was queried first in a trial was counterbalanced within subject.

Adults were told that they were to match dimensions. For example, on the size–darkness trials, one triplet of objects was placed on the table in ascending or descending order (e.g., three cubes varying in size) and then the second unordered set was placed on the table (e.g., three cubes varying in darkness). The subject was asked "Which grey is like the first size, which is like the second, and which is like the third. Line up these cubes so they are in the same order as these." On the size and darkness trials, the dimension serving as the demonstration dimension was counterbalanced within subjects. On the size–loudness and darkness–loudness trials, the sounds served as the demonstration dimension. Subjects were asked to match the cubes to a series of sounds played on the tape recorder. On half the trials of each type the sounds were ascending in loudness and on half they were descending. Subjects were asked "Which size (grey) is like the first sound, which is like the second, which is like the third. Line up the cubes here so they are in the same order as the sounds. Put the cube that is like the first sound here (on the left), the one that is like the middle sound here, and the one that is like the third sound here (on the right)." The sounds were replayed as often as the subject requested.

In both the children's and the adults' task there were a total of 12 trials resulting from four replications of each of the three pairwise combinations of size, darkness, and loudness. These 12 trials were embedded in a larger task of 24 total trials that included cross-dimension matches between three other combinations of dimensions that will not be reported here. The order of the 24 trials was randomly determined for each individual.

Results and Discussion

Figure 8 shows the numbers of individuals at each age level making 0, 1, 2, 3, or 4 of the designated matches for each dimensional combination. (The frequency of the inverses of the designated matches is given by 4 minus the number of designated matches.)

Performances in the Size–loudness and Size–darkness conditions replicate the first two experiments. The number of individuals making big–loud/little–quiet matches four out of four times increases reliably with age, $\chi^2_{(2)} = 27.75, p < .001$. The number of subjects making big–dark/little–light matches four out of the four possible times changes reliably with age, $\chi^2_{(2)} = 9.65, p < .01$ with 4 to 5 year olds less likely to make those matches than 2 to 3 year olds and adults. Developmental changes are also apparent in the third new condition in which subjects matched loudnesses to darknesses. The increase with age in subjects making dark–loud/light–quiet matches on all four trials does not reach statistical significance, $\chi^2_{(2)} = 3.78$. The increase with age in individuals making the opposing dark–quiet/light–loud matches on at least three of the four trials approaches statistical significance, $\chi^2_{(2)} = 5.1, p < .06$. Nonetheless, the

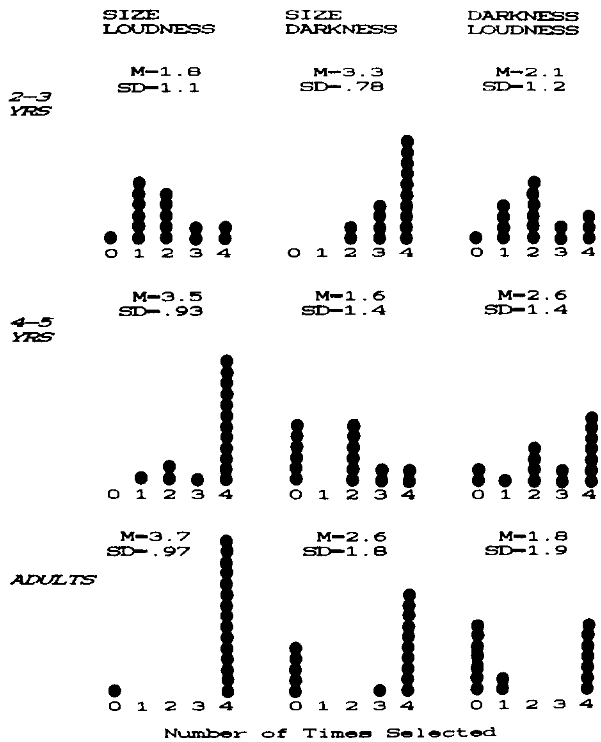

FIG. 8. Number of individuals making big–loud/little–light matches, dark–loud/light–quiet matches, big–dark/little–quiet matches 0, 1, 2, 3, or 4 of 4 possible times. Means (M) and standard deviations (SD) are given at the top of each frequency distribution.

developmental trend is clear. Early, there is little uniformity in darkness–loudness maps—within or between individuals. By adulthood, both possible maps are made; for some adults in this task, dark is like quiet and for others, dark is like loud.

The central question of this experiment is the relation between cross-dimension maps: If an individual maintains that big is like loud and loud is like dark, does that individual judge big to be like dark? To answer this question, an individual's judgments must be consistent. Accordingly, for this analysis, we included only subjects who mapped *all* three pairs of dimensions in one direction three or four out of four times. By this criterion, only four 2 to 3 year olds, eight 4 to 5 year olds, and all sixteen adults were consistent. The reason so few 2 to 3 year olds contribute data is apparent in Fig. 8; many fail to meet the consistency requirement on size–loudness and/or darkness–loudness trials.

Of the 28 subjects who were consistent, we asked whether their cross-dimension mappings were transitive. There are a total of 8 sets of pairwise

TABLE 2
Number of Subjects at Each Age Level Who Made Consistent Cross-Dimension Maps of the Eight Possible Kinds

	2–3 years	4–5 years	Adults
Transitive			
Big = Dark = Loud	—	3	7
Big = Light = Loud	—	—	5
Big = Dark = Quiet	—	—	2
Big = Light = Quiet	—	—	—
Total transitive	0	3	14
Not transitive			
Big = Dark/Big = Loud/Loud = Light	4	2	2
Big = Quiet/Big = Dark/Loud = Dark	—	—	—
Big = Loud/Big = Light/Loud = Dark	—	3	—
Big = Quiet/Big = Light/Loud = Light	—	—	—
Total not transitive	4	5	2
Total consistent	4	8	16

mappings across the three dimensions; 4 of these are transitive and 4 are not. Table 2 lists the 8 possible sets of mappings and the numbers of subjects who made each. Of the four 2 and 3 year olds who made consistent pairwise mappings, all *violated* transitivity. All four of these children consistently judged big to be like loud, big to be like dark grey, but loud to be like light grey. Of the eight consistent 4 to 5 year olds, only three children's judgments were transitive. In contrast, 14 of the 16 adults' judgments were transitive. The two sets of transitive mappings that dominate adult judgments (12 of the 14 transitive sets of judgments) are the ones in which (1) big is like loud is dark (7 individuals) and (2) big is like loud is like light (5 individuals). In brief, it would seem that adults organize dark as like big and loud *or* as like little and quiet.

The results of Experiment 2 suggested that the *mature* organization of dimensions is constrained such that perceptual structure and lexical structure agree within in an individual. The results of Experiment 3 suggest further that, in the mature system, the organization of an individual dimension is made to be consistent with the organization of other dimensions. If big is like loud, and loud is like dark, then within the mature dimensional system, big must also be like dark.

One might argue that the transitivity of adult judgments is spurious. Individual adults might not have any systematic organization of the three dimensions. Instead, adults might simply respond to implicit task demands—making ad hoc mappings of the dimensions but doing so consistently from trial to trial. The fact that subjects' 12 mappings across these three-dimensional combinations were randomly intermixed among 12 trials asking them to map other dimensions argues against an explicitly strategic approach to this task. Nonetheless, we sought further evidence on this issue by asking 14 adults who were unfamiliar with this research to make a different cross-dimension judgment (size–darkness, size–loudness, darkness–loudness) on each of three days that were separated by more than four intervening days. The subjects, faculty, staff, and graduate students in the psychology building were casually approached by a different experimenter on each of 3 days and asked their "opinion" of how two dimensions were related. Specifically, with no stimuli in view, the subjects were asked by different experimenters on different days (1) whether big was like loud or like quiet, (2) whether big was like dark grey or light grey, and (3) whether loud was like dark grey or light grey. The hope was that the separation in time, context, and experiment would minimize attempts by subjects to be consistent across the three judgments. Despite these efforts, 11 of the 14 individuals' cross-dimension maps were transitive. This preliminary evidence suggests individual adults' organizations of cross-dimension similarities are stable and internally consistent.

EXPERIMENT 4

For adults, it seems reasonable to suppose that big is like loud *because* big is more than little and loud is more than quiet. What is the basis of the youngest children's cross-dimension matches? Is big like dark grey for young children *because* big is more than little and dark grey is more than light grey? What is the basis of individual adults' matches of darkness and size? Do some adults believe that dark is more than light whereas others believe that light is more than dark? Experiment 4 addressed these questions by directly asking children and adults which end of a dimension was the "more" end. We examined the dimensions of size, darkness, loudness, and hue.

We included hue in order to gain greater insight into young children's perceptual organization of achromatic color. One possible account of children's treatment of dark grey as like big is that surface darkness reminds young children of nighttime and big scary things. Another possible account is that darker grey is more than lighter grey because of children's experiences with coloring. Children start with white paper and add color to it. White may therefore be the zero point in children's organization of color. Both of these accounts are problematic in that it is unclear why nighttime fears or coloring should organize 2 year olds' judgments so strongly but not 3 and 4 year olds.

A third, and we think more likely, possibility, is that young children's organization of achromatic color is a symptom of the sensory structure of the color space as a whole. This possibility is suggested by the way some languages divide up the color space (see Berlin & Kay, 1969). Languages with two or three color terms lexically organize the color space into a dark region that includes the darker greys, black, and the colors of (focal) blue, (focal) green, and purple, and into a light region that includes the lighter greys, white, yellow, orange, and sometimes red. There is also evidence suggesting that children learning English naturally segregate colors into two categories of dark (or "cool") and light (or "warm," see Park, Tsukagoshi, & Landau, 1985). The question we asked, then, is whether the polar organization evident in young children's perception of achromatic color is also evident in their perception of chromatic colors. Is dark grey more than light grey and is green more than orange for very young children?

Method

Subjects. The subjects were twelve 2 year olds (2;4 to 2;11), twelve 4 year olds (4;0 to 4;6), and twelve undergraduates. There were equal males and females in each group and all subjects were tested in the laboratory at the Psychology Department. Three additional 2 year olds were tested and replaced because they failed to meet the criterion for understanding the task as outlined below. Ten additional 2 year olds were tested in a control procedure.

Stimuli. Three pairs of stimuli were constructed for each of five dimensions: (1) *Number*—(a) 25 Popscicle sticks bound together versus 4 Popscicle sticks bound together, (b) a plate of 10 toy candies versus a plate of 2 toy candies, (c) a 2.5 × 3.5-cm card with 32 stickers on it versus a 2.5 × 3.5-cm card with 5 stickers on it; (2) *Size*—(a) a 30-cm-tall cylinder versus a 7.5-cm-tall cylinder of the same width as the first, (b) a 20-cm circle versus a 5-cm circle, (c) a 37-cm-long and 20-cm-high pick-up truck versus a 7.5-cm-long and 5-cm-high pick-up truck; (3) *Darkness*—all pairs consisted of 6 × 9-cm pieces of grey paper that varied in Coloraid notation from 1 (almost white) to 9 (black), (a) 1 versus 5, (b) 3 versus 7, (c) 4 versus 9; (4) *Hue*—all pairs consisted of 6 × 9-cm hue sheets, in Coloraid notation, (a) R(ed) versus B(lue), (b) O(range) versus G(reen), (c) Y(ellow) versus BVB(purple); (5) *Loudness*—2 s of clapping at 55 dB SPL versus 80 dB SPL, (b) a 1.4-s 610-Hz tone at 66 dB SPL versus 82 dB SPL, (c) a 1.0-s 900-Hz whistle at 66 dB SPL versus 80 dB SPL. The sounds were recorded and played on a Marantz tape recorder.

Procedure. The session began with training trials which used the Number stimuli. On these trials, the child was introduced to a toy bear. The child was told "The bear always wants the one that is more. He (she) always wants the one that is a lot." The first number pair was displayed on the table and the child was asked "Which is more? Which is a lot? Give the bear the one that is more." If the child did not select correctly, feedback as to the correct selection was provided. The child was given two passes through the three number contrasts to choose correctly three times in a row. Three 2 year olds did not meet this criterion and were replaced. Four additional 2 year olds were corrected at least once. All remaining children responded perfectly on the first three trials. After meeting criterion on the Number trials, the next 12 trials were presented in one of two random orders with the dimension of difference intermixed. On each trial involving visual stimuli, the objects were placed on the table with the right–left position of objects within a pair randomly determined. The tape-recorded auditory stimuli were played by depressing a button. The order of the louder and quieter sounds alternated across trials (but not within trials). On each trial involving visual stimuli, the child was asked "Which is more?" Which is a lot? Give the bear the one that is more." On the loudness trials, the two sounds were played and then the child was asked "The bear likes the one that is more. Which sound do you think the bear likes? Which one is more? Which one is a lot? This one (the first sound is replayed) or this one (the second sound is played). Listen again, is this one more?" The first sound was played and the experimenter waited for a yes–no answer. Then the second was played and the experimenter asked, "Is this one more?" and waited for a yes–no answer. If on any trial (visual or auditory), a subject hesitated or said there wasn't one that was more (or that both were), the subject was asked which one the bear would think was more. Subjects were forced to choose on each trial.

Adult subjects were told that this experiment was being conducted with very young children and that for comparison purposes we needed to ask them the very same questions in the very same way that we asked them of children. The adult subject was then introduced to the bear and the procedure was identical to that used with children.

Ten additional 2 year olds were tested in the identical procedure except that the words *more* and *a lot* were not used. On all trials, the children were asked "Pick one for the bear. Which one do you think the bear likes? Give the bear one of these." These trials examined the possibility that young children do not understand the intensive direction signaled by the words *more* or *a lot* but simply pick the positive pole item. Choices under this procedure did not deviate from chance for any dimension, $t(9) < 1.48$, in all cases, and will not be considered further.

Results and Discussion

The frequency distributions of children's and adults' selections of the bigger object, the darker grey, the darker color (blue, green, purple), and the louder sound as *more* are shown in Fig. 9. A 3(Age) × 4 (Dimension analysis of variance of these data revealed a main effect of Dimension, $F(3,99) = 11.85$, $p < .001$, and a reliable interaction between Age and Dimension, $F(6,99) = 3.34$, $p < .01$.

Post hoc analyses (Tukey's B, $\alpha = .05$) indicated no reliable age differences in the choice of the bigger object as more. There was almost perfect performance in these judgments at each age level. However, there was a reliable *decrease* between ages 2 and 4 years in the choice of the darker grey as more and between 2 year olds and both 4 year olds and adults in the choice of the darker hues as more. Finally, at each successive age level, subjects were reliably more likely to select the louder sound over the quieter one.

The results for the 2 year olds fit well with the findings of the first three experiments. Size and darkness, the dimensions that have consistently designated more ends for 2 year olds, are the same dimensions across which young children make uniform cross-dimension maps. The fit between the present data and those of the first three experiments may seem less good for the 4 year olds and adults. Specifically some 4 year olds in this experiment had difficulty in picking the louder of two objects as "more" whereas in the previous experiments children at this age readily mapped loud sounds to big objects. One possible reason for this is that the 4 year olds in this experiment are young, all are below 4½ years. Another

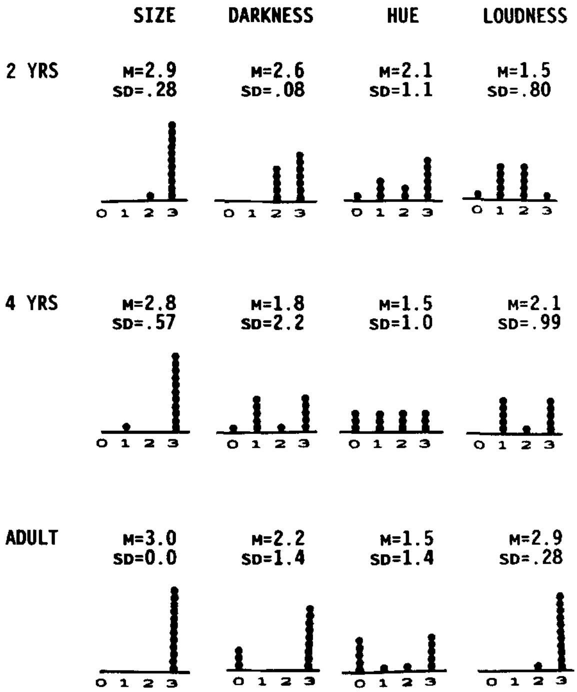

FIG. 9. Number of individual's choosing the big cube, the dark grey, the dark hues, and the loud sound as more. Means (M) and standard deviations (SD) are given at the top of each frequency distribution.

possibility is that loud sounds at this age are more strongly linked to big percepts and the word *big* than to the words *more* and *a lot*.

Also, in the present experiment, 9 out of the 12 adults designated dark as more than light whereas in Experiments 2 and 3 just slightly more than half the adults made the analogous map of dark to big. We suspect that this is a sampling effect. Some adults maintain that dark is more than light (and therefore like big). Others maintain that light is more than dark (and therefore like big). In this experiment we may simply have had more adults of the first kind than the second. Note, however, the general character of the trend in this study is consistent with the previous findings: For size, there is little developmental change. For loudness, there is much between-subject variability early in development but none late in development. For darkness there is little variability early in development, much in the middle, and then sharp individual differences late in development.

Finally, the finding that young children polarize hue in the same direction that they polarize achromatic color suggests that their early magnitude marking of achromatic color reflects something basic about the perceptual structure of color. We calculated the Pearson *r* coefficient between the number of choices of the dark grey as "more" and the number of choices of the dark hues as "more" separately for each age level. These coefficients are .88 for 2 to 3 year olds, .65 for 4 to 5 year olds, and .18 for adults. Very young children's patterns of hue and achromatic color choices are strongly related whereas adults' choices across the two dimensions of color are not. Thus, it would seem that with development, achromatic and chromatic color become differentiated.

The next experiment compared children's and adults' designation of the "more" ends of dimensions and dimension words.

EXPERIMENT 5

Method

Subjects. Sixteen subjects at each of three age levels participated: 2 to 3 year olds (2;4 to 3;3), 4 to 5 year olds (4;8 to 5;9), and undergraduates. The children were tested at a local daycare. The undergraduates were tested in the laboratory of the Psychology Department.

Stimuli. The stimuli in the *Percept* task were identical to those used in Experiment 4 with the exception that the hue stimuli were not used. The stimuli in the *Word* task consisted of the word pairs *big–little, dark–light, loud–quiet.* These same dimensional terms were queried in a Word comprehension task.

Procedure. All children completed the experiment in two sessions—a Percept session and a Word session on separate days with less than 10 days intervening. The adults completed the experiment in one session. The order of the Percept task and Word task was counterbalanced across subjects at each age level. Each session began with a training procedure that was identical to the procedure used in Experiment 4. All children tested met the criterion for continuing. The Percept task was identical to that of Experiment 4 except each of the unique trials was repeated twice for a total of 18 trials. The order of these 18 trials was randomly determined. In the Word task, the subject was told that the Experimenter was going to the store to buy a present for the bear. The subject was told "I can buy bear a ______toy or a ______toy. Which one is more? Is ______more or is ______more"? Each pair or opposing dimensional terms was queried six times in this way. The order of mention of "more" and "less" terms in the carrier phrases alternated between trials in a randomly determined manner.

Only the children participated in the Word-comprehension task. This task was conducted at the end of the Word session. The child was presented twice with each of the three unique Size-contrast stimuli and was asked on one trial for each contrast to indicate the *big* one and on the other trial to indicate the *little* one. Analogously, the child was presented twice with each of the three unique Loudness contrasts and on separate trials for each contrast indicated the *loud* sound and the *quiet* sound, and was also presented twice with each of the three Darkness contrasts and indicated the *dark* paper and the *light* paper on separate trials. These 18 word-comprehension trials were intermixed and presented in a randomly determined order.

Results and Discussion

Figure 10 shows the mean choices of the bigger, louder, and darker objects in the Percept and Word task and Fig. 11 shows scatterplots of

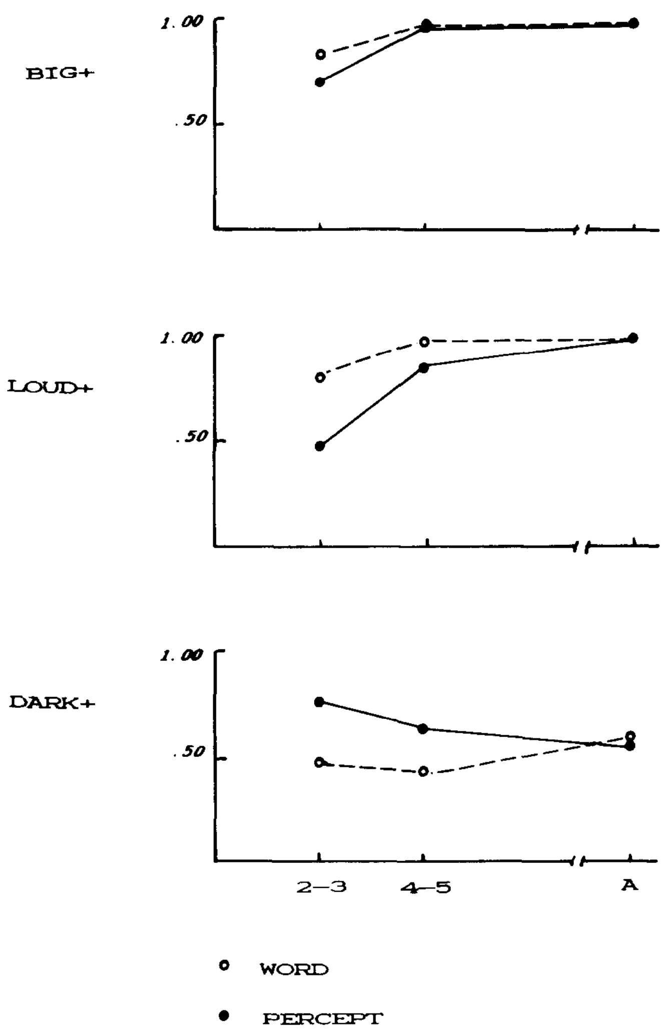

FIG. 10. Mean proportion choices of big, loud, and dark as "more" in the Percept and Word tasks of Experiment 5.

each individual's choices in the Word task as a function of their choices in the Percept task. Table 3 gives the mean performances of children in the Word-comprehension task. Children who did *not* understand both of the queried terms perfectly (that is, six out of six correct choices for each dimension in the Word comprehension task) are indicated by open figures in the scatterplots of Fig. 11. We consider performances on each dimension separately.

Size. Each subject's numbers of choices of the bigger object and the

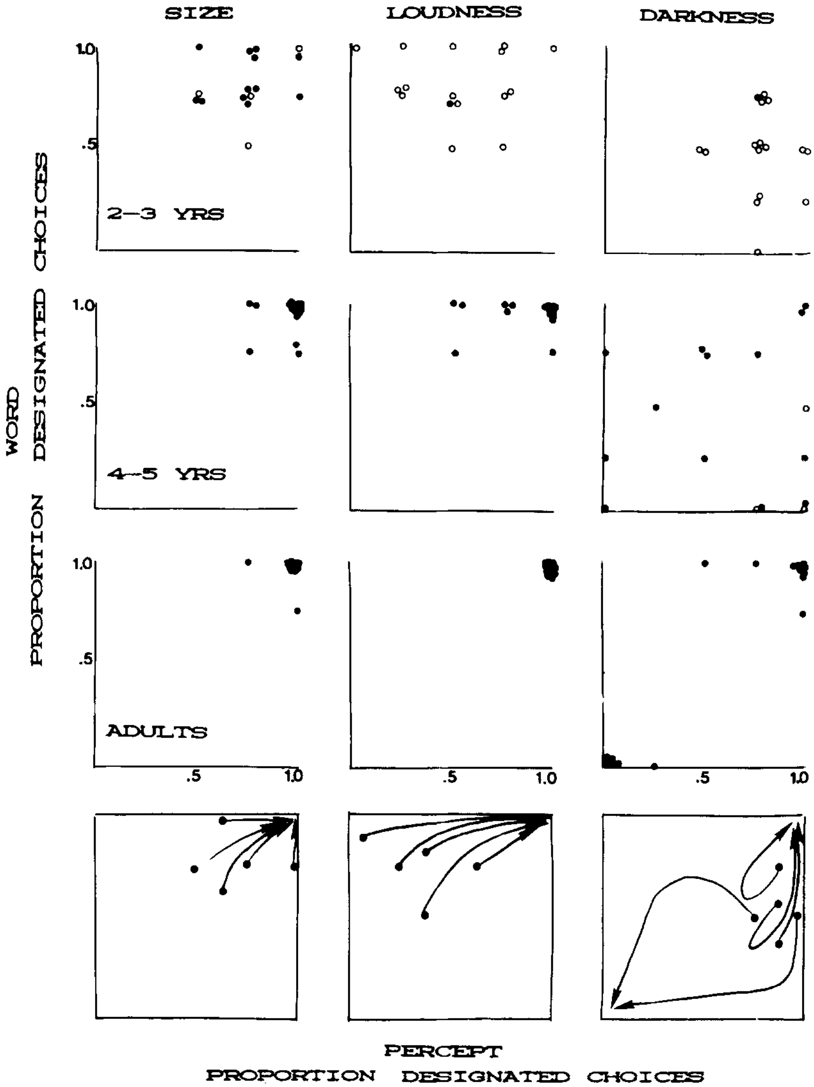

FIG. 11. Scatterplots of individual's proportion choices of big, loud, and dark as "more" in the Percept and Word tasks of Experiment 5, and in the bottom panels theoretical developmental trajectories in the same space, from left to right, big is more, loud is more, and dark is more.

word *big* were submitted to an analysis of variance for a 3 (Age) × 2 (Task) × 2 (Order) mixed design. The analysis revealed a reliable main effect of age, $F(2,42) = 38.82$, $p < .001$, and a reliable Age × Task interaction, $F(2,42) = 3.89$, $p < .05$. Older children virtually always chose

TABLE 3
Mean Proportion Correct Responses in the Word–Comprehension Task of Experiment 5

	Big	Little	Loud	Quiet	Dark	Light
2–3 years old	.96	.94	.67	.67	.54	.62
	(14)	(13)	(8)	(8)	(3)	(5)
4–5 years old	.98	1.00	1.00	1.00	.92	.92
	(15)	(16)	(16)	(16)	(13)	(13)

Note. Number of children (max = 16) choosing correctly on 3 out of 3 trials is given in parentheses.

the bigger object and the word *big*. The youngest children also chose the bigger object and the word *big* but less consistently than the older subjects. Moreover, the youngest children were more likely to choose the word *big* over the word *little* than to choose the bigger object over the little one. Notice also that virtually all the children, including the youngest ones, showed near-perfect comprehension of the words *big* and *little* in the Word-comprehension task. Moreover, despite the reliable developmental differences and the task effect for the youngest children, it is quite clear in the scatterplots of Fig. 11 that even the youngest children have good knowledge that big is more than little and that the word *big* signals more than the word *little*.

Loudness. The developmental trend for loudness is similar to that for big only more pronounced. The 3 (Age) × 2 (Task) × 2 (Order) analysis of variance revealed reliable main effects of Age, $F\ (2,42) = 49.18$, $p < .001$; Task, $F\ (1,42) = 21.95$, $p < .001$; and a reliable Age × Task interaction, $F\ (2,42) = 9.60$, $p < .01$. The average performance of the youngest children in the Percept task is at chance ($t\ (15) < 1.00$) but is reliably better than chance in the Word task ($t\ (15) = 20.81$, $p < .001$). The advantage of the word *loud* in indicating more for young children over the perception of a loud object is obvious in the individual data as well as in the group means. Among the youngest children, there is less variability in the Word task (the dots indicating individual subjects are only in the top half of the state space) than in the Percept task (the dots indicating individual subjects vary over the entire horizontal range).

The youngest children's performances in the Word task suggests that these children know that the word *loud* signals the same intensive direction as the word *more*. However, many of these same children showed imperfect comprehension of the meanings of *loud* and *quiet* in the Word-comprehension task. Apparently, children's early partial knowledge of the meaning of *loud* includes its relation to the word *more* before it specifies how *loud* applies to the relative amplitude of sounds.

At maturity, loudness possesses a "more" end both linguistically and perceptually. Older children and adults pick—with considerable uniformity—the louder object as well the word *loud* as more. The older children also show near-perfect comprehension of the words *loud* and *quiet*.

Darkness. The developmental trend with darkness is quite unlike that with size of loudness. The 3 (Age) × 2 (Task) × 2 (Order) analysis of variance of the number of dark-is-more responses revealed only a main effect of Task, $F(1,42) = 6.29$, $p < .02$; darker objects were chosen as "more" with greater frequency than the word *dark* was designated as "more." The lack of reliable differences between age groups in this analysis stems directly from high between-subjects variability at the older age levels.

The scatterplot of individual performances shows 2 to 3 year olds to mostly fall above .50 in their choices of dark as more in the Percept task; thus again, the youngest subjects maintain that darker greys are more than lighter greys in the Percept task [$t(15) = 24.18$, $p < .01$]. Their performances vary about chance and as a group do not reliably differ from chance [$t(15) = 1.16$] in their choice of the word *dark* or *light* as indicating *more*. This chance level performance is not surprising since few of the youngest children understood the meaning of *dark* and *light* as applied to achromatic color.

The older children, in contrast, show a much better understanding of the words in the Word-comprehension task but as a group show no systematic choice of the darker or lighter object as *more* or of the word *dark* or *light* as indicating *more*. Moreover, eight of the sixteen 4 and 5 year olds show discordant patterns of performance in the Percept and Word tasks. Seven children systematically maintain that the darker paper is *more* in the Percept task but choose the word *light* as indicating *more* in the Word task. At the time that children understand the words *dark* and *light* as they apply to achromatic color, their magnitude marking of the dimension and the words may be unrelated.

Finally adults' performances are highly organized. Some individual adults maintain that dark is more. Some maintain that light is more. But whatever choice an individual adult makes in the perception task, the same choice is made in the word task.

The strength and frequency of 4 to 5 year olds' and adults' choices in this experiment fit well with the results of Experiments 1–3 and clarify the results of Experiment 4. That is, 4 to 5 year olds judge loud to be more than quiet and to be like big. Adults divide as to whether dark is more or light is more and they divide as to whether dark or light is like big. Thus, it would seem that both the cross-dimension matching task and the "Which is more" task are tapping the same developmental phenomena.

Developmental trajectories. The results of Experiment 5 are summarized by the theoretical trajectories depicted at the bottom of Fig. 11. The data for size and loudness suggest a uniform and highly constrained developmental pathway. Big is perceptually and linguistically more than little early in development and it stays that way. Loud starts out more disorganized and young children know that the word *loud* means more than *quiet* before they reliably know that loud sounds are more than quiet ones. But as in the case of size, the polar organization of loudness moves in the same direction in perception as it does in language and the developmental movement is rapid. Darkness, in contrast, shows an irregular development trend. Early, dark is perceptually more than light. But the acquisition of language co-occurs with the disruption of this early polar organization. Marked individual variability and within individual discordant percepts and words appear. At maturity, the individual differences remain but percepts and words agree.

Before proceeding to our discussion of these findings, we report one last experiment that examined 2 year olds' apparent lack of knowledge that loud sounds are more than quiet ones. This knowledge seems so secure in our adult conceptions that it is difficult to accept the young child's failures. Moreover, the temporal nature of sound and the memory demands it placed on successful performance raises questions about whether our various tasks were sensitive measures of 2 year olds' optimal performance. Accordingly, in a long series of pilot studies, we tried many procedures (including training) to demonstrate that 2 year olds did know that loud was more than quiet. None of these procedures was successful except one. We report that success as Experiment 6.

EXPERIMENT 6

Method

Subjects. The subjects were ten 2 year olds (2;1 to 2;9), five males and five females.

Stimuli and procedure. Two toy police cars, one 5.6 cm long and one 29 cm long and two toy 4×4 trucks, one 8.7 cm and one 35 cm long served as the size stimuli. The sounds made by the two larger vehicles (when activated by a switch) were recorded on a Marantz tape recorder. In the experiment, the siren sound was played at 57 dB and 82 dB SPL and the truck motor sound was played at 55 dB and 86 dB.

At the beginning of the task, the child was shown a toy cow and a toy dog and told that one of them had made a sound on the tape recorder. The sound (moo) was played and the child was asked which animal made it. All children answered correctly. The experimental trials were then begun. On each trial a sound was played and the child indicated the object making the sound. There were eight trials. On half, the sound was loud; on half, the sound was quiet. On half of the loud trials and half of the quiet trials, the sound was a siren and the child chose between police cars. On the remaining trials, the sound was the motor sound and the child chose between 4×4 trucks. The eight trials were presented in one of two randomly determined orders.

Results and Discussion

The children chose the large vehicle on 60% of the trials when a loud sound was played and on 40% of the trials on which a quiet sound was played. The difference is statistically reliable, $t(9) = 3.00, p < .05$. Further, the five boys who participated chose more correctly than did the girls, 68% versus 52% correct. The results indicate that in a well-defined context, 2 year olds (and particularly male 2 year olds) know that big vehicles make louder sounds than little ones.

This result, of course, need not mean these 2 year olds know that big is more than little and that loud is more than quiet. The children may only know that bigger trucks make louder sounds and that bigger police cars have louder sirens. They may not know anything general; they may not know that bigger and louder are "more than" directions on their respective dimensions. The fact that 2 year olds (and particularly 2 year old boys) show this knowledge in a real-world task that builds on the real-world correlation between the size and loudness of vehicles suggests that the more general correspondence between size and loudness that we see at maturity may be partially built on the empirical correspondences that exist in the world.

GENERAL DISCUSSION

The principal results from the six experiments are these: (1) The magnitude marking of the poles of quantitative dimensions changes with development and the developmental pattern is different for different dimensions. Experiments 4 and 5 show that the big end of size is stably "more" throughout development, whereas the behavioral designation of the loud end of loudness as "more" becomes stronger with development, and the attribution of a "more" end to darkness changes with development.

(2) Cross-dimension similarities change with development. Experiments 1–3 indicate unidirectional growth in the similarity of sizes and loudnesses. With development, big becomes like loud and little becomes like quiet. In contrast, the developmental changes in the similarities of sizes and darknesses are not unidirectional. Big is like dark and little is like light early in development but this cross-dimension similarity becomes variable and may even reverse for some individuals with increasing age.

(3) The acquisition of dimension words occurs at the same time as the changes in perception. Experiments 2 and 5 indicate that comprehension of dimensional terms is temporally related to changes in cross-dimension similarities and in the magnitude marking of dimension poles. The perceived similarity between big and loud and the ability to pick the louder object as "more" happen at the same developmental time as children understand the words *loud* and *quiet*. The dissolution of the perceived similarity between big and dark and the increased variability in the designation of the more end of darkness happen at the same time that chil-

dren begin to understand the words *dark* and *light* as they apply to achromatic color.

(4) Amidst the variable developmental histories of specific dimensions, there is a general trend toward a unified perceptual and lexical organization. Experiments 2 and 5 indicate that, with development, the "more" and "less" ends of perceptual dimensions and the "more" and "less" meanings of dimension words come into agreement.

(5) The mature system of dimensions appears to be internally consistent. Experiment 3 showed that adults' cross-dimension correspondences are transitive—a fact that implies a unified system of dimensions. Coupled with the finding that words and percepts *come into agreement* with development, this result suggests that from the inchoate dimensional structures of children, a unified, coherent, *logical* system of dimensions may emerge. Using Gentner and Rattermann's (1990) term, perceptual dimensions, cross-dimension similarities, and dimension words may become gentrified with development—neatened up and made orderly.

(6) The dynamics of development are different for different dimensions. Size and loudness develop rapidly to apparently universal endpoints. The developmental dynamics of darkness are more complicated; there appears to be a single starting point but then irregular and multiple paths to several endpoints.

What do these results tell us about the origins of the psychological structure of dimensions? We consider three possible accounts: (1) The polar structure of dimensions is principally determined by the sensory physiology. (2) The polar structure of some dimensions (size and loudness) is determined by the sensory physiology but the sensory physiology leaves other dimensions (darkness) unconstrained and for these unconstrained dimensions language plays an influential role. (3) The structure of all dimensions emerges in development from the interaction of sensory, perceptual, cognitive, and linguistic factors.

The Polar Structure of Dimensions Reflects the Sensory Structure

In light of the present results, can we retain the idea that the more and less ends of quantitative dimensions are determined by the sensory physiology? The central problem in retaining this idea is determining whose data reflect the sensory physiology—the 2 year old or the adult? One possibility is that they both do and that unknown but critical sensory neural structures mature in a way that causes dissolution of 2 year olds' judgments that dark is more than light. This is an unlikely possibility. Current evidence indicates that the human sensory system is virtually mature by the end of the first year (see Aslin & Smith, 1988 for a review).

A second possibility is that adults' performances but not 2 year olds' reflect the underlying sensory structure. One might assume as Stevens did that the adult psychophysics reflect the *given* structure of dimensions—that size and loudness are innately quantitative dimensions and darkness is not. By this view, the reason size and loudness develop rapidly to one endpoint is that their structure is determined by the underlying physiology, and the reason darkness is characterized by variable development and outcome is that there is no innate constraint on its structure.

The problem with this possibility is that we must find some way to dismiss the developmental data; we must find some way of explaining away the fact that 2 year olds consistently judge dark grey to be more than light grey whereas 4 year olds do not. We might attribute the developmental differences to performance variables or task demands—to factors that somehow obscure the true sensory structure of the dimensions. This account is neither promising nor satisfying. It is not promising because there are no known performance variables that readily explain why 2 year olds as a group are *more* consistent in their judgments of darkness than are 4 year olds but are at the same time and in the same tasks *less* consistent than 4 year olds in their judgments of size and loudness. This approach is not theoretically satisfying because it presumes that mature performance is a better guide to the origins of dimensional structure than is the developmental path to the mature structure.

The third possibility is that the 2 year olds performances but not the adults' performances reflect what is given in the sensory system. This possibility makes sense; task strategies, conceptual structures, language, and other cognitive processes are more likely to *systematically* intervene between the sensory structure and adults' performances than between the sensory structure and the performances of 2 year olds. The stature and contributions of Stevens notwithstanding, the present evidence suggests that darkness is originally and perhaps most fundamentally a quantitative dimension with darker the direction of increase. If this conclusion is correct and if there is no difference in the sensory processing of darkness for 2 year olds and adults, the implications are profound for it would mean that the sensory structure is neither the sole nor final determiner of the *perceived* structure of a dimension.

The Polar Structure of Dimensions Reflects Underlying Sensory Structure or Language.

A modified Stevens position is worthy of consideration. Some dimensions, namely the prothetic dimensions, may be specified by a sensory physiology and maintain the same structure throughout development. The perceived structure of prothetic dimensions such as size and loudness may be relatively impenetrable by language and other learning. The sensory structure of metathetic dimensions such as darkness, however, may (despite the consistent judgments of 2 year olds) be more complex, allow for multiple organizations, and thus be influenceable by language.

In her discussion of nouns and verbs, Gentner (1982) argued that the relevant language-thought question was not whether linguistic determinism or perceptual-cognitive determinism is correct, but the proper question is *which* applies *where*. She argued that in some cases the perceptual system tightly constrains a single outcome so that language has no choice but to follow the perceptual-cognitive system's lead. In other cases, the perceptual system may allow for multiple organizations that language selects among. The idea, extended to the present case, is that the perceptual structure of size and loudness is tightly constrained by the sensory structure and is uninfluenceable by language and other learning. In contrast, the sensory structure of darkness, while initially pointing to dark as more than light, allows for other organizations as well and thus the perceptual organization can be altered by language. The evidence on human color perception and language fits this proposal in that the human color space affords multiple organizations only some of which are selected by any language.

The human color space is a double cone in which perceived dissimilarity is represented by distance. Figure 12 illustrates some of the possible organizations of this space: A shows the focal hues; B shows the "diagonal" split of the color solid into the two broad categories of dark and light; C shows the partition of the color solid into the chromatic colors and the achromatic core of the desaturated colors (white, greys, and black). In some languages, lexical categories are principally organized around the focal hues. In other languages, there are only two or three color terms organized (like the perceptual judgments of 2 year olds) around the dark–light split of the space (see Berlin & Kay, 1969; Bornstein, 1973; Heider & Oliver, 1972).

Could language be a causal factor, then, in the organization (or loss of organization) of the polar structure of darkness? The present evidence is certainly not conclusive. However, the temporal relation between the changes in perception and language are suggestive. Moreover, the direction of development is *toward* the structure of language. *Dark* and *light* are not unambiguously more and less terms in English and the direction of development is away from an initial organization that clearly marks dark as more toward a structure that is (at least across individuals) ambiguous. Perhaps, then, the acquisition of the system of semantic markedness of English adjectives (Greenberg, 1966) works to weaken the initial perceptual organization of dark as more than light.

There is a second property of English which may also work against a dark-is-more organization. English uses the same words to talk about surface darkness and illumination. We use *dark* to refer to colors such as

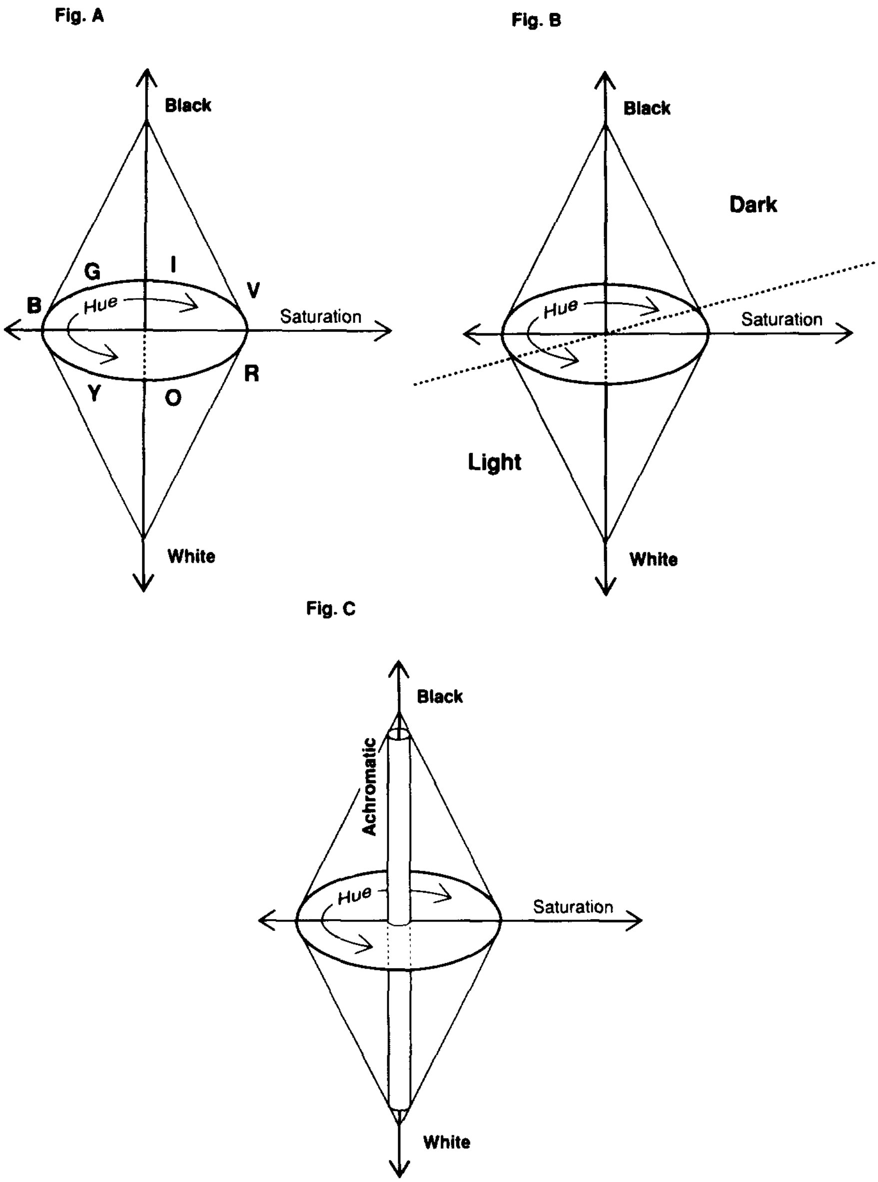

FIG. 12. (A) Coordinates of the human color space. Darkness is represented by the vertical dimension. The hues are arranged in a circle with saturation maximum at the outer edges. (B) A partition of the space into dark and light. (C) A partition of the space into the chromatic and achromatic.

navy blue and black and to refer to a lack of illumination. We use *light* to refer to pink and white and as the name for illumination itself. English thus wrongly encourages us to think of the achromatic core of the color solid and intensive variation in illumination as the same dimension with the same "more" and "less" ends. This alignment of achromatic color with brightness may also be encouraged by the perceptual relation between surface color and illumination at the *extremes* of illumination; that is, when there is no light, the perceived color of all objects is dark grey and when there is intense light, the perceived color is blinding white.

Two testable hypotheses are suggested by the idea that English selects among possible organizations of color: First, if 2 year olds' judgments reflect the underlying sensory order, and *if perceived* structures but not sensory processing can be altered by language, then there ought to be tasks that tap the sensory processing at very early levels that will show dark to be more than light for adults. Treisman and Gormican (1988) have reported evidence that is consistent with this idea. Second, if 2 year olds' judgments reflect a perceptual structure that may or may not be selected by language, then there ought to be some languages in which dark is more than light. Languages that merit investigation in this regard are those that have different words for both illumination and surface darkness (for example, Hebrew) and those like some West African languages that lexically categorize the color space into dark–light regions and thus do not by their lexical categories segregate the achromatic from the chromatic colors.

The second account explains the developmental trend for darkness by positing competing and interacting forces. Language influences perception and the developmental trend is irregular because color allows for multiple organizations. But by this second account, there is only one possible perceptual organization of size and loudness as evidenced by the uniform, directed, developmental trend to a single outcome. We believe this second account is promising in its explanation of darkness but fails in its explanation of the origins of the polar structure of size and loudness. This second account fails as did the first because it requires a dismissal of some of the developmental data—in this case, the developmental changes in children's designation of loud as "more" and like big. These judgments become more consistent at the same time that judgments of darkness becomes less consistent. On what grounds should we dismiss these data? Are the data from adults and our adult theoretical belief that the polar structure of size and loudness is "given" more important in deciding between developmental hypotheses than the changing performances of children as they mature?

We propose a third account that explains the developmental trajectories for loudness and size in the same way we explain the development of darkness. We suggest that the apparently complex interactions that exist between perception and language in the case of darkness exist for all dimensions. The odd case of darkness reveals a more general truth about developmental process.

A Systems Account

Specifically, we suggest that the polar structure of all dimensions emerges *in* development *in* the interaction of three factors: the sensory system, cross-dimension relations in language, and cross-dimension relations in the world. This systems approach explains the variable developmental path and variable outcome for darkness in terms of multiple *antagonistic* forces. The sensory organization, language, and cross-dimension correlations in the world (for example, between brightness and darkness at the extremes of illumination) all pull in different directions and probably with different strengths for different individuals. No coalition of these multiple forces is universally strong enough for a single developmental path; however, small variations in any one may be sufficient to send an individual's development in a particular direction.

We explain the uniform and seemingly constrained developmental pathways for size and loudness in terms of multiple *converging* forces. We suggest that 2 year olds have greater difficulty in mapping sizes to loudnesses than sizes to darknesses because of the inherent dissimilarity of sights and sounds. This early *real* lack of a perceived similarity between big and loud does not matter for long because many other forces push for one alignment of the two dimensions. Language designates both *big* and *loud* as "more" terms. Our use of *big* as a general scalar in constructions such as *big noises* explicitly relates big and loud. The physical structure of the world also supports an organization in which big is like loud and little is like quiet. Size and loudness are correlated, albeit imperfectly: Bigger objects tend to make more noise than littler objects. Experiment 6 suggests that very young children possess knowledge of this correlation and that it helps them relate big and loud. By this view, language and correlations in the world push loud to be like big just as language and correlations in the world push dark to become less like big.

In brief, we suggest that size and loudness develop rapidly to a single outcome not because the sensory structure predetermined it to be so but because several converging forces make it so. By this view, darkness, size, and loudness differ not so much in their sensory organizations or in what is "innate" as in the language used to talk about the dimensions and in each dimension's relations to other dimensions in the world. These forces converge so tightly in the cases of size and loudness that it is easy to imagine that the outcome is predetermined by some hardwired component. However, there may be no predetermined epigenesis (Gottlieb, 1991). Instead, the mechanism of development may require and expect the experiences of language and dimensional correlations. When all these forces agree, as will usually be the case, we have an overdetermined endpoint with rapid and uniform developmental change. When the forces compete, we set the stage for variability in the developmental trajectory and developmental outcome.

We find this systems view compelling. Unlike the alternative accounts, the systems approach provides a framework in which we can hope to explain all the data and need not dismiss some findings as irrelevant. The account is parsimonious; one explanation is given for the development of all dimensional structures. The account makes evolutionary and developmental sense. We have evolved *to develop* adaptive cognitive structures; we are not given predetermined structures that may not fit where we find ourselves.

A systems account may also help us explain the global direction of development. We found that growth within individuals is toward an internally consistent set of dimensional organizations that are transitive and in which language and perceptual structure agree. Stability and systematicity are the natural products of a system propelled forward by multiply interacting and mutually constraining forces. A discordant organization of perceptual dimensions and dimension words (or nontransitive cross-dimension maps) is inherently unstable; changes in task and context will push the system one way at one moment and another at the next. Systematicity is the ultimate developmental outcome because a stable state may not be reached until systematicity occurs. This proposal affirms classic notions of cognitive development (see Piaget, 1963 and also recent discussions by Fentriss, 1984; Oyama, 1985; Thelen, 1989).

CONCLUSION

We began this paper by asking whether the structure of dimensions is fixed by the perceptual system or whether dimensional structures can be influenced by language. Our results indicate that dimensional structures are not fixed and may be shaped by language. However, in light of the data, we may want to rephrase the question. Moreover, the field may benefit from a broader revision than Gentner's (1982) rephrasing of the question as "*which* applies *where*." The proper question seems not to be one of whether language in any domain "leads" or "follows." The interrelations between language and perceptual-cognitive structure seem to be of a different nature, and include the possibilities of facilitation, attenuation, and reciprocal influence. Simply, there may be no "yes" or "no" answer to the question of whether language determines perception or perception determines language for *any* lexical category at any level of analysis. The proper question is not *whether,* or *which,* or *where* but only *how* perception and language mutually influence each other.

Our point and the value of a *developmental* systems approach to cognition is clarified by a metaphor offered by van Geert (1991). He views development as like the evolution and colonization of an island biotope. Perception and perceptual language can be thought of as two species in this biotope. The adaptations of each species clearly depends on each other and all other species on the island. No adaptations can be understood in isolation. Moreover, it makes no sense to ask whether one species *determines* the adaptation of the other. The outcome of development, the structure of the island biotope as a whole and the adaptations of the individual species, is best understood as a dynamic system of continual interaction and mutual influence.

By what methods may we come to understand such a developmental process in which multiple parts mutually and continuously influence each other? As a first step, we suggest a shift from the study of knowledge states to the empirical study of the developmental pathway itself—its points of stability, instability, organization, and reorganization. A topological description of the shape of development may provide useful insights to the various forces on development—when they converge and when they compete.

REFERENCES

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution.* Berkeley: Univ. of California Press.

Bierswisch, M. (1970). On semantics. In J. Lyons (Ed.), *New horizons in linguistics* (pp. 164–184). London: Penguin.

Bierswisch, M., & Lang, E. (1989). *Dimensional adjectives: Grammatical structure and conceptual interpretation.* New York: Springer-Verlag.

Blewitt, P. (1982). Word meaning acquisition in young children: A review of theory and research. In W. Reese (Ed.), *Advances in child development and behavior* (Vol. 17, pp. 139–195). San Diego: Academic Press.

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

Bornstein, M. H. (1973). Color vision and color naming: A psychological hypothesis of cultural difference. *Psychological Bulletin,* **80,** 257–285.

Bornstein, M. H. (1974). The influence of visual perception on culture. *American Anthropologists,* 77, 774–798.

Carey, S. (1982). Semantic development: The state of the art. In G. Wanner and L. R. Gleitman (Eds.), *Language acquisition: The state of the art.* London/New York: Cambridge Univ. Press.

Clark, H. H. (1973). Space, time, semantics, and the child. In T. Moore (Ed.), *Cognitive development and the acquisition of language.* New York: Academic Press.

Fentriss, S. C. (1984). The development of coordination. *Journal of Motor Behavior,* **16,** 99–134.

Gentner, D., & Rattermann, M. J. (1990). Language and the career of similarity. In S. Gelman (Ed.), *Language and conceptual development.*

Glucksberg, S. (1988). Language and thought. In R. Sternberg & E. E. Smith (Eds.), *The psychology of thought* (pp. 214–241). New York: Cambridge Univ. Press.

Gottlieb, G. (1991). Epigenetic systems view of human development. *Developmental Psychology,* **27**, 33–34.

Greenberg, J. H. (1966). *Language universals.* The Hague: Mouton Publishers.

Heider, E. R. (1971). "Focal" color areas and the development of color names. *Developmental Psychology,* **4**, 447–455.

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

Holyoak, K. (1978). Comparative judgments with numerical reference points. *Cognitive Psychology,* **10**, 203–243.

Johnston, J. R. (1985). Cognitive prerequisites: Cross linguistic study of language acquisition. In D. Slobin (Ed.), *Universals of language acquisition: Vol 2. Theoretical issues* (pp. 961–1004). Hillsdale, NJ: Erlbaum.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: Univ. of Chicago Press.

Levin, I. (1982). The nature and development of time concepts in children: The effects of interfering cues. In Freedman (Ed.), *The developmental psychology of time* (pp. 47–85). New York: Academic Press.

Lucy, J. A., & Shweder, R. A. (1979). Whorf and his critics: Linguistic and nonlinguistic influences on color memory. *American Anthropologist,* **8**, 581–613.

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch and loudness. *American Journal of Psychology,* **87**, 173–188.

Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities.* New York: Academic 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.

Oyama, S. (1985). *The ontogeny of information: Developmental systems and evolution.* Cambridge, MA: Cambridge Univ. Press.

Park, Tsukagoshi, & Landau, B. (1985). *Young children's mis-naming of colors.* Presented at the biennial meeting of the Society for Research in Child Development, Toronto, Canada.

Piaget, J. (1963). *The origins of intelligence.* New York: Norton.

Smith, L. B. (1985). Young children's attention to global magnitude: Evidence from classification tasks. *Journal of Experimental Child Psychology,* **39**, 471–492.

Smith, L. B. (1987). Perceptual relations and perceptual language: Commentary on Marks et al. *Monographs of the Society for Research in Child Development,* **52**, (1).

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, MA: Cambridge Univ. Press.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review,* **64**, 153–181.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects.* New York: Wiley.

Thelen, E. (1989). Self-organization in developmental processes: Can systems approaches work? In M. Gunnar (Ed.), *Systems in development: The Minnesota Symposium in Child Psychology* (Vol. 22). Hillsdale, NJ: Erlbaum.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review,* **95,** 15–48.

van Geert, P. (1991). A dynamical systems model of cognitive growth: Competition and support under limited resource conditions. For L. Smith & E. Thelen, *A dynamical systems approach to action and cognition: Applications*. Cambridge, MA: MIT Press, manuscript in preparation.

(Accepted May 28, 1991)