In C. Rovee-Collier (Ed.) (1998), Advances in infancy research, Vol. 12 (pp. 355-373). Norwood, NJ: Ablex Publishing Co.

Introducing the 3-D Intermodal Preferential Looking Paradigm: A New Method to Answer an Age-Old Question

George J. Hollich and Kathy Hirsh-Pasek
Temple University

Roberta Michnick Golinkoff
University of Delaware

Introducing the 3-D Intermodal Preferential Looking Paradigm: A New Method to Answer an Age-Old Question

Amy, mother of 4-year-old Joshua and 10-month-old Allison, is helping her 4-year-old find his ball. She says, "Where did you see the ball last?" At that moment, Allison crawls purposely to the couch and digs the ball out from the corner between the cushions where it is barely visible! Amy is shocked. Clearly, Allison understood the word "ball!

This drama, repeated in countless forms and countless houses across the world, aptly illustrates a paradox of developmental psychology. While parents are certain that their children know more than they say, until recently, psychologists did not have the methodological tools to confirm this. The past two decades, however, have witnessed a kind of mini-revolution in experimental procedure with the development of paradigms that allow researchers to peer into infants' formerly hidden competencies. Developmental psychologists now have at their disposal a number of tools and methods to examine early infant cognition and behavior. Using looking, sucking, and surprise as dependent measures, age-old questions about nascent abilities for language are being revisited.

One such age-old question concerns the origins of word learning. Between the ages of 17- and 24- months, infants' vocabulary literally explodes, with infants learning, on average, seven to nine new words a day (Carey, 1978; Carey & Bartlett, 1978). What happens in the months preceding this vocabulary spurt to precipitate such a dramatic eruption of linguistic brilliance? How do these children sift through the myriad of potential meanings to learn words in their language so quickly and apparently, so effortlessly? To answer these questions, two prominent positions have developed. Principles or constraints theories posit that built-in principles serve to constrain the word learning situation. In contrast, social/pragmatic theories place their emphasis on the interaction between the apprentice and the experienced word learner. Whatever the perspective, the debate seems to revolve around the competence of the very youngest word learners (see Hirsh-Pasek & Golinkoff, 1993).

This paper introduces the Three Dimensional (3-D) Intermodal Preferential Looking Paradigm, a method specifically designed to study questions of the origins of word learning. Using real, 3-D objects, preferential looking, and a highly controlled (yet realistic) labeling situation, the 3-D procedure allows for careful examination of the developing processes by which pre-vocabulary spurt infants come to map labels to objects.

In the first portion of this paper, the 3-D paradigm is introduced. In the next two parts, some preliminary results garnered from the first large scale use of this procedure are presented. In the final section, the 3-D paradigm is evaluated. We consider the ages for which it is appropriate, the kinds of tasks for which it is best suited, and the possibility of using it to study infant language comprehension from a hybrid theoretical orientation. That is, we take an approach that embraces both the constraints and social pragmatic theories.

The 3-D Intermodal Paradigm

Brief History

There is a long line of methods to study children's emerging vocabulary and grammatical abilities. These include pointing tasks, picture tasks, surveys of mothers' knowledge, head turn procedures, preferential listening procedures, and preferential looking procedures (McDaniel, McKee, & Cairns, 1996, for a review). In keeping with this tradition, the 3-D Intermodal

Preferential Looking Paradigm has its roots in work done by Baldwin (in press) and Fagan (1971; Fagan, Singer, Montie, & Shepard, 1986). Principally, however, the 3-D paradigm is a modification of the intermodal preferential looking paradigm developed by Hirsh-Pasek and Golinkoff (1993; 1996a; 1996b; Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987),

The Intermodal Preferential Looking Paradigm (see Hirsh-Pasek & Golinkoff, 1996a; 1996b; or Naigles, this volume) was adapted from work by Spelke (1979), who developed it to study intermodal perception. In it, the infant is seated on a blindfolded parent's lap in front of two laterally spaced video monitors. A concealed centrally placed audio speaker plays a linguistic stimulus that matches only one of the displays shown on the screens. A hidden observer records the total amount of time (measured to hundredths of seconds) that the infant spends watching the matching versus the non-matching screen (see Figure 1, or Figure X in Naigles, this volume).

To see how this works, consider the case of noun comprehension (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). A center light directs the child's attention towards the center, and a linguistic message (produced in child-directed speech) is delivered from the concealed speaker. The child hears, "where's the APPLE?" On one screen, the child sees a BELL and on the other an APPLE, and the linguistic message is "Do you see the APPLE? Show me the APPLE!"

The logic of this procedure, which has been consistently confirmed, is that children look more quickly and longer at the screen displaying the targeted object (the APPLE) than at the screen displaying the non-targeted object (the BELL). That is, infants give more attention to the video event that matches what they are hearing, in this case the linguistic message, than to a video event that does not match. Indeed, in no studies performed in our laboratories have children ever shown a significant preference for the non-matching screen. In a way, this seems reasonable. Parents often comment on the actions, objects, and events surrounding the child, clearly intending to direct the child’s attention to the next object, action or event. A large portion of the adaptive advantage conferred by language depends on orienting in a manner consonant with the language message.

The major advantages of the classic “2-D” paradigm are numerous. First, only a looking response is required from the child. Thus, children do not have to point, answer questions, or act out commands. This permits testing at very young ages, even as young as four and a half months (Spelke, 1979). Second, the video monitors can represent dynamic stimuli. This is critically important for the study of verbs, as well as complex syntactic frames. Third, children are able to take advantage of a coalition of syntactic, semantic, and prosodic information: all on-line. That is, language learning occurs amidst a coalition of cues which are rarely, if ever, separated in the actual input to the child. In the real world, unlike experimental designs, children do not experience semantic probability divorced from syntactic complexity. Thus, competencies which were previously suspected might be revealed if, and only if, a large coalition of cues is available to the child.

For these reasons, the 2-D Intermodal Preferential Looking Paradigm has proven very useful in the study of early language. Some of the results thus far have indicated that infants appear sensitive to cues for constituent structure by 14 months of age (Hirsh-Pasek & Golinkoff, 1996), to the meaning of common nouns and verbs by 16 months of age, (Hirsh-Pasek & Golinkoff, 1996) to word order by 17 months of age (Hirsh-Pasek & Golinkoff, 1996), to morphological word endings by about 18 months of age (Golinkoff, Hirsh-Pasek, and Schweisguth, 1997), and to information about sentence frames by 24 months of age (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Hirsh-Pasek & Golinkoff, 1991; 1993; Naigles, 1990; this volume; Bavin, this volume; see Hirsh-Pasek and Golinkoff, 1996b for a review). In each case, the paradigm has proven a sensitive measure of the cues that infants and toddlers use to extract information from the input, particularly for very young children.

Like any method, however, there are a few limitations. For example, as with any infant study, data collection is only possible for as long as the attention span of the child lasts. Children sometimes lose interest in "playing the game," and depending on the age, some of data can be lost in this manner. Yet, in the studies above, and in many ongoing studies, the data is both consistent and robust. Also, the paradigm is a highly technical procedure: synchronizing videos, lights, cameras, and computers can be a daunting task and is one that requires a solid support staff.

More centrally, however, the 2-D paradigm is limited in that it can not examine the social interaction between adults and children, nor does it permit the child to explore real objects or otherwise act on the scene before them. Many believe that these interactions are critical for the word learning process, particularly at early ages (Baldwin, in press; Tomasello & Farrar, 1984). As a result, it would seem appropriate to expand upon the initial 2-D procedure in ways that would permit the study of early word learning with higher levels of social interaction. Thus, the 3-D intermodal preferential looking paradigm, which incorporates real, 3-D, objects and a live experimenter was developed to increase ecological validity while at the same time maintaining strict experimental control. Somewhat unintentionally, this increased emphasis on social interaction had the added benefit of significantly reducing subject loss, particularly for the youngest ages (11-12 months).

Overview of the 3-D Paradigm

The 3-D Intermodal Preferential Looking Paradigm accomplishes functionally similar objectives to the 2-D paradigm by means of an experimenter visible to the baby, a video camera, and a special display board. The design of the display board was taken from an apparatus constructed by Fagan (1971) in his infant test of attention. In the 3-D paradigm, infants are presented with real 3-D objects that are labeled in a play environment by the experimenter. These same objects are then presented side by side on the Fagan board during test trials while the now hidden experimenter requests one or another of the objects. As in all such preferential looking paradigms, the infant's task is simply to watch the displays. Any preference for one object over another will be detected by subsequent scoring of the video record. The premise is the same as in the 2-D paradigm. It is expected that the children will look at that stimulus that matches the linguistic message they are hearing. While this method does not allow for the use of dynamic stimuli, it nonetheless proves very engaging for children by having real objects and a live experimenter.

Apparatus

The basic setup is portrayed in Figure 2. The infant is seated on a blindfolded parent's lap 75 cm back from the center of the modified Fagan box which sits on a table. This modified box consists of a 55 cm x 50 cm base, and a hinged 40 cm x 50 cm board (see Figure 3). The board is painted black on one side with velcro attachments at 20 cm from the top and 12.5 cm from either side (providing two sites for attaching objects 30 cm apart). The board is hinged such that it can rotate lengthwise, pivoting to hide or reveal whatever objects might be attached, and thereby providing precise control over the duration of exposure. Timing of the presentation is accomplished by use of a specially designed timer (Infant Test Timer) which can be set to produce a brief tone after the requisite period has elapsed. A mirror behind the infant allows the video camera to record not only the infant looking responses (which are directly visible) but also the objects on the board and the behavior of the experimenter (which are both visible through the mirror). Observers (blind to the condition being run) make use of the specially designed timer to calculate, from the video record, infant's total looking times to each object. As in previous preferential looking studies, inter- and intra-rater reliability is kept above an r value of .90. This is accomplished by training raters to a 90% correlation criteria and is further validated by random re-codings of one third of the data. The data from any rater is not permitted unless all scores are greater than 90% in agreement. Typically, the mean correlation is higher (mean r =.96).

General Procedure

The general procedure consists of alternating familiar and novel trials. In the familiar trials, children are tested on their comprehension abilities for the names of familiar objects. These familiar trials provide an introduction to the paradigm itself and are composed of two phases: exploration and testing. Initially, they were also incorporated to assess if children were capable of performing in the paradigm. Conversely, the novel trials (in which two novel objects are presented and one is labeled), are used to assess nascent labeling abilities in very young children and are composed of four phases: exploration, salience, labeling, and testing.

In the exploration phase, the child is presented, sequentially, with each toy to play with for 26 seconds. (The order of presentation is counterbalanced). This gives the child a chance to have a full range of haptic experience with the objects before the remaining phases and provides time in which the experimenter engages the child's interest and lays the groundwork for subsequent social interaction.

In the salience phase, the experimenter hides centrally behind the board. The experimenter turns the board to reveal the two objects side by side,and subsequently presents a neutral linguistic stimulus (eg: "Look up here! What do you see?"). After six seconds, the experimenter turns the board back (hiding the objects) and stands in preparation for the next phase. Infant looking times to each object provide a baseline measure of the relative salience. If the visual stimuli have been well balanced for perceptual factors, attention (in the form of looking times) should be evenly distributed across the objects. If the stimuli are intentionally unbalanced on perceptual factors, looking times should reflect this. The experimenter is hidden throughout the salience phase in order to limit the possibility of biasing the child's response. The neutral linguistic stimulus is used because results from pilot studies indicated this further engaged children in the task.

In the labeling phase, the child sees two objects presented on the table while one of the objects is labeled (in some manner) by the experimenter. Usually, this involves getting the child's attention by saying the child's name and following with something like, "Horace, look at the MODI -- the MODI. It's a MODI." However, one advantage of this paradigm is its flexibility, and the range of possible labeling phases is nearly as diverse as the range of actual labeling situations in the real world.

In the testing phase, the experimenter again hides centrally behind the board, out of view of the infant, and the child again sees the objects presented on the display board for six seconds. This time, however, before hiding or turning the board, the experimenter further engages the child via the linguistic query, "Where's the X?" (where X is the name of the targeted object). In pilot studies, this question proved critical to focus children on the task. Only after this verbal query does the experimenter hide and turn the board to reveal the objects. While the objects are visible, the child hears a linguistic message asking, "Do you see the X? Look at the X." After six seconds, the board turns back over, hiding the objects. The test phase is then repeated in order to obtain a more reliable assessment of the child's looking times.

As in previous preferential looking paradigms, the dependent variable is mean looking time to each target in both the labeling phase and in the testing phase. The independent variables change from experiment to experiment but usually include the age of subject, gender, and one or two principle experimental manipulations (e.g. direction of experimenter's eye gaze, direction of pointing, perceptual salience, temporal contiguity, etc.). Variables which are counterbalanced include order of presentation, side of presentation, and the objects used.

The next two sections recount some of the preliminary results garnered from the first large scale use of the 3-D paradigm. The outcome from the familiar trials of this study are presented in the next section while the findings from the novel trials are presented in the subsequent section. This separation of results is done principally because each set of trails has a radically different rationale, addresses different questions, and produces data which are best analyzed and discussed separately. Further, a thorough review of these results is beyond the scope of this paper and will be published elsewhere (Hollich, Hirsh-Pasek, & Golinkoff, in preparation). Here, we seek merely to demonstrate the validity of the paradigm and to suggest, via example, some of its potential uses.

The Familiar Trials: A Validation of the 3-D Method

Rationale

Children are notoriously difficult subjects. Getting consistent, usable data is often extremely complicated and time consuming. This is particularly true in the study of infants, since short attention spans make it difficult for them to stay on task for more than a few minutes. Given this limitation, the advent of any new method raises questions about the usefulness and ability of that method to extract data in a robust and timely manner.

In particular, before one can assess (via the 3-D paradigm) whether or not children have learned a new label, one has to verify that the paradigm is even capable of detecting word knowledge in the first place. If the 3-D Paradigm is to be a valid measure of labeling abilities, then subjects should look longer at the appropriate familiar object when it is requested than the inappropriate one (just as familiar, yet not requested). If children do not look longer at requested familiar objects (objects whose labels they know), there is no reason to posit that they should do so for requested novel objects which have been recently labeled. If children (even pre-vocabulary spurt infants) do indeed show a longer looking response toward requested familiar objects, then the 3-D paradigm would seem a valid measure of infant language knowledge.

Method

Subjects were tested at two different locations: the Temple University Infant Laboratory and the University of Delaware's Infant Language Project. They were recruited either from published birth records (Delaware laboratory) or from a purchased list (Temple laboratory). A total of 110 children were tested in this study, but the data from 29 of the children had to be discarded due to experimental error, parental "peeking," fussiness, or a side bias (greater than 75% attention to one side) on the part of the children. Thus, 81 children served as subjects unevenly distributed across three age/vocabulary ranges: 12 to 13 months/pre-vocabulary spurt (n = 26; 14 males, 13 females; mean age = 12.58 months, productive vocabulary = 8.63 of the words listed on the MacArthur Communicative Development Inventory, infant version), 19 to 20 months/mid-vocabulary spurt (n = 30; 13 males, 17 females; mean age = 19.46 months, productive vocabulary = 128.57 of the words listed in the MacArthur CDI, toddler version), and 24 to 25 months/post vocabulary spurt (n = 25; 17 males, 15 females; mean = 24.27 months, productive vocabulary = 304.63 words on the MacArthur CDI, toddler version). Stimuli were two sets of paired "familiar" objects: ball/book and keys/block.

Upon arrival at the laboratories, children and their parents were invited into a playroom. While the child acclimated to the laboratory environment, the purpose of the task was explained to the parents. The child and the parents felt ready, they moved into the testing room. Once the child was seated comfortably on a parent's lap, the parent put on a visor to cover his or her eyes and testing began.

The familiar trials consisted of an exploration phase and testing phase (see Figure 4). In the exploration phase, the child was allowed to play with each toy, individually, for 26 seconds. (The order of presentation of these toys was counterbalanced). In the testing phase, the experimenter got the child's attention by calling the child's name and saying, "Where's the X?" (where X is the targeted familiar object). The experimenter then crouched behind the display board, the display board was rotated so that the objects were visible (for six seconds), and the experimenter asked, "Do you see the X? Show me the X?" Then the board was rotated to hide the objects, and the experimenter stood, again called the child's name, and repeated, "Where's the X?" The experimenter again crouched behind the board, and the board was rotated to reveal the objects (for six seconds), and again the experimenter asked, "Do you see the X? Show me the X?" Note that while the experimenter was visible during the <u>exploration</u> phase and was permitted to interact and engage the child at the beginning and middle of the testing phase, throughout the actual <u>testing</u> the experimenter was not seen. This was done to minimize the possibility of the experimenter biasing the child's looking times. Additionally, the targeting was counterbalanced such that each object was requested half of the time, and used as the distractor the other half.

<u>Results and Discussion</u>

Mean percentage of looking times (with standard deviations in parentheses) to the targeted object were .55 (.16), .63 (.15), and .66 (.13) for the 11-12 month, 19-20 month and 24-25 month groups, respectively (see Figure 5). Equal looking times would be reflected by a percentage score of .50. However, with an alpha level of .05, pre-vocabulary spurt (12-13 months), mid-vocabulary spurt (19-20 months) and post vocabulary spurt children (24-25 months) looked significantly longer at the targeted than at the non-targeted object (all $\underline{t}$'s $< .05$). Thus, it appears that even 11-12 months old children are capable of the preferential looking response necessary for this task.

Two other points about this validation study are of interest. First, not only is this method successful for 12 month olds, but there is very little subject loss. Across the entire study, only five of the 110 subjects were lost due to fussiness. (The reasons for the other 24 losses are listed under subjects.) Second, children distribute their looking times between the targeted and non-targeted objects and do not look exclusively to one object.

In sum, it appears that there is evidence for a robust method that potentially allows examination of hypotheses about the very earliest word learning. The next section illustrates one application of the 3-D Intermodal Preferential Looking Paradigm while expanding on the theoretical perspective that we take for the development of both word learning and mature grammar: the Hirsh-Pasek and Golinkoff (1996b) Coalition Model (Golinkoff & Hirsh-Pasek, 1995).

The Novel Trials: A Preliminary Study of the Cues to Reference

Rationale

To address the question of how children learn words so rapidly, psychologists have proposed a number of apparently opposed classes of theories, two of which are of particular relevance to this discussion. One class of theories, "constraints" or "principles" theories, solve the induction problem by positing that children entertain only certain hypotheses for a word meaning. In essence, the child comes to the task of word learning prepared with certain principles or strategies for how to map words onto referents. For example, perhaps children innately assume that words (though not music) refer to an object, action or event and that words map onto whole objects rather than onto object parts (Markman, 1987, 1992; Golinkoff, Mervis, Hirsh-Pasek, 1994). A second class of theories, constructivist or social pragmatic theories, place the solution to the word learning problem in the social interaction between the apprentice (i.e. the child) and the sophisticated conversational partner (i.e. the parent). In this view, the interaction with the environment is the critical element (Bloom, 1993; Tomasello & Akhtar, 1995).

In general, these theories have had a polarizing effect on the word learning literature (Tucker, Hirsh-Pasek, & Hollich, in press; Hirsh-Pasek & Golinkoff, 1996b), with the constraints camp focusing on driving principles, and the constructivists on environmental interaction. However, Golinkoff, Mervis, and Hirsh-Pasek (1994) have developed a coalition view of the word learning conundrum, thereby adopting what has been called “the radical middle.” Perhaps, mature principles of the principles/constraints theories are the *products* of development, not the *engines* of development. For example, children may not begin with a full blown principle like reference (that words refer to objects). Rather, this principle might emerge on a continuum as children acquire their very first words. Indeed, reference may begin as simple associative learning (a “goes with” kind of representation). Children might associate a word with the first visible object they see, relying on temporal contiguity, or pehaps they might associate the word with the most perceptually salient object they see. At some later age, around 19 to 20 months of age, children would then abandon this associative principle in favor of one that is more sophisticated -- a 'stands for' relationship between word and referent (Baldwin, in press). Eighteen month-old children, for example will follow the eye gaze of the speaker to determine word reference as if trying to interpret the speaker's intent over and above the environmental lures (Baldwin, 1991, 1993, in press). Twelve-month-old children may not. By incorporating a developmental perspective, this model takes a midline position between a constraints approach and the constructivist position, positing that children’s lexical development is the product of complex epigenetic interactions between weighted constraints and the external social environment.

In the Hirsh-Pasek and Golinkoff (1991; 1996b; Golinkoff & Hirsh-Pasek, 1995) framework, children are assumed to be surrounded by a coalition of multiple input sources (see Figure 6). As development progresses, principles develop which predispose children to be sensitive to certain aspects of that input and not to others, depending on the task. These developing principles provide the groundwork that allows children to engage in a type of “guided distributional learning” in which they attend to certain cues in the input and use these to construct mature word learning principles (Hirsh-Pasek and Golinkoff, 1996b). Thus, mature principles are emergent phenomena. Importantly, multiple sources are always available at the outset of language learning, and the differences that are seen throughout development result because children are differentially sensitive to some of these cues over others at different points in development.

Among the coalition of cues that children use to form these heuristics or principles for word learning are perceptual cues (movement, shape, texture), functional cues (what the object does), temporal cues (which objects they see first, or which object are temporally contiguous with labeling) syntactic cues (word order, inflection, etc.) and social cues. Thus, the social pragmatic view is no longer inherently inconsistent with the principles account of word learning. Rather, it is an integral part of the foundation for the construction of those principles. While still endorsing a principles-type account, this newer examination of the constraints allows for an integrative approach to the theories in the field incorporating both the data from the social pragmatist and the lexical principles camp.

Further, the coalition view and the theories it incorporates were introduced to illustrate that through the 3-D paradigm, the idea that word learning is the product of developing sensitivities is now an empirically testable one. The 3-D preferential looking paradigm permits a examination of the coalition theory, cue by cue. Indeed, because the 3-D procedure can obtain data for pre-vocabulary spurt (11- and 12-month-old) subjects and because it permits interactive use of social cues, it seems to provide an optimal procedure to systematically vary various social, perceptual, functional, and temporal cues in a controlled labeling situation. Thus, children's weighting and use of these cues at various ages can now be established, and the relative contributions of constraints and social interaction can be empirically validated.

To give a more concrete example and introduce our preliminary study, earlier work by Baldwin (1991, 1993, in press; Baldwin & Moses, in press), suggests that by 19 to 21 months, children can use <u>social eye gaze</u> to override cues like temporal contiguity when learning a label. However, it has been noted that children's attention is also strongly affected by <u>perceptual</u> cues like color and complexity (Golinkoff, Mervis, & Hirsh-Pasek, 1994). Indeed, we suggested above that the principle of reference may go from simple association with perceptually salient objects to a more mature principle which weights social factors, like eye gaze. For the novel trials of the preliminary study, then, the cues of social eye gaze were pitted against perceptual salience. That is, instead of keeping the stimuli equally interesting, one of the objects to be labeled was designed to be much more interesting than the other. In one condition, the <u>coincidental</u> condition, the experimenter labeled the interesting toy using only social eye gaze (not handling or pointing to it). Thus, both the perceptual cues and social eye gaze worked together (in coincidence) to establish a label. In the <u>conflict</u> condition, the boring toy was labeled so that social eye gaze and perceptual salience were in conflict with one another (see Figure 7). In this manner then, by comparing infants' behavior in each of the conditions, one could examine the changing weighting of social eye gaze and perceptual salience over time.

We hypothesized that very young children, with an immature principle of reference, would be drawn by perceptual cues when attaching a word to a referent; that the word might "go with" the most interesting perceptual object in the environment. We further hypothesized that older children of 19 and 24 months of age with a more mature principle of reference might not only follow subtle social cues like eye gaze, but would use those social cues to assume that a label refers to an object, even when those social cues were in conflict with perceptual cues. Thus, they would "read" the speaker's intent when affixing a label to a referent, even when the object to be labeled is the more boring alternative in the environment.

<u>Method</u>

Subjects were the same as those reported in the familiar trials above. However, in this case the subjects were further divided into two conditions, the coincidental and conflict conditions. For the conflict condition n = 15, 14, 18, in the 11-12, 19-20, 24-25 month-old groups, respectively. For the coincidental condition n = 12, 16, 14, for the 11-12, 19-20, 24-25 month-old groups, respectively (see Figure 7).

The procedure used for the novel trials, as set out in the general procedure described above, consisted of exploration, salience, training, and testing phases. An expanded view of the novel trials is presented in Figure 8. In the exploration phase, the child played with each toy individually for 26 seconds. In the salience phase, children heard a neutral linguistic stimulus (like "Look up here! What do you see?") and saw the boring and interesting object presented side by side on the display board for six seconds. In the labeling phase, the child saw both objects presented on the table, while one of the objects was labeled by an experimenter. Which object was labeled was determined by the condition (coincident or coincidental). The experimenter first got the child's attention by saying the child's name, locking eye gaze, looking down at one of the objects, and saying, "look at the X -- the X." (X is the novel word to be used as a label.) Next, the experimenter again got the child's attention by saying the child's name, and again, after locking eye gaze, attempted to direct the child's attention to one of the objects while saying again, "look at the X - the X. It's a X." In this case, the novel words used were MODI and DANU, for the first and second pairs of novel objects respectively.

In the testing phase, the children again saw the objects presented on the display board for 6 seconds. This time, however, they first heard the linguistic message, "Where's the X?" Next, the experimenter hid, and the board turned to reveal the objects. While the objects were visible, the children heard a linguistic message asking them, "Do you see the X? Look at the X." After six seconds had elapsed, the board turned back over to hide the objects from the infant's sight. The test phase was then repeated.

Results and Discussion

Preliminary results for the salience, training and testing portions of the study are summarized in Figures 9 and 10. Although incomplete, there are several points worth noting. First, examination of the salience scores suggests that children at all ages did attend more to the interesting toy than the boring toy when social cues and/or labeling were not provided (during the salience trials). Thus, in the absense of social cues, it does indeed appear that across all three age groups, children looked at the "interesting" toy much longer than at the "boring" toy. (The mean percentage of looking time [with standard deviations in the parentheses] to the interesting toy for the 11-12, 19-20 and 24-25 month-old groups, respectively, were .66 [.16], .61 [.13], and .64 [.10].) With an alpha level of .05, the effect of salience was statistically significant at $p < .0001$ for all of the groups ($t(26) = 5.25$, $t(29) = 4.82$, $t(31) = 7.93$, for the 11-12, 19-20 and 24-25 month-old groups, respectively). Thus, we suitably selected stimuli which the children found appropriately boring or interesting.

Second, it appears that social eye gaze can make a difference in the training phase for the older age groups. That is, mean percentage of looking times to the interesting toy were significantly different ($p < .0001$) in the conflict versus the coincidental condition for the 19-20 and 24-25 month-old groups (with an alpha level of .05, $t(28) = 8.48$, and $t(30) = 7.37$ for the 19-20 and 24-25 month-old groups, respectively). Thus, despite salience differences, when a toy was labeled (with accompanying social eye gaze), 19-20 and 24-25 month olds looked significantly longer at it than when that toy was not labeled. Conversely, eye gaze alone (in conjunction with labeling) seemed to have no effect during the training phase for younger infants (11-12 months old).

Finally, the same pattern of results appears in the testing trials. That is, mean percentage of looking times to the interesting toy were, again, significantly different ($p < .001$) in the conflict versus the coincidental condition for the 19-20 and 24-25 month-old groups (with an alpha level of .05, t (28) =_4.73, and t (30) = 4.09 for the 19-20 and 24-25 month-old groups respectively). Thus, when a previously labeled toy was requested, 19-20 and 24-25 month olds looked significantly longer at it than when the other toy was labeled previously and requested. Not suprisingly (not having been able to follow the eye gaze in the training phase) younger children (11-12 month-old) did not appear to look significantly longer at the requested object. Notice, however, that in the testing trials, the mean percentage looking times to the interesting toy do not drop significantly below 50% (which would indicate a preference for the boring toy). Given the strong salience preference, this is not particularly surprising but might bear further examination in subsequent studies.

Nonetheless, the data the we have presented are consistent with the coalition model in postulating the late emergence of social compared with perceptual cues. Only at 19-20 months of age do children appear make use of social eye gaze in a word learning situation. This data are also consistent with evidence from a large social pragmatic literature. Tomasello and Todd (1983) find that mothers who follow their children's focus of attention (especially in early infancy) tend to have children with higher vocabularies (see also Tomasello & Kruger, 1992; Rehill, Heberle, Hirsh-Pasek & Golinkoff, 1996). Mothers who attempt to lead their children to their own focus of attention are not as successful until the children reach about two years of age. While the data are tantalizing, even at this preliminary stage, a detailed discussion and analysis of these data would be both premature and beyond the scope of this paper. These admittedly preliminary results are given simply to suggest the usefulness of the 3-D paradigm in empirically testing how various cues to word learning are weighted at different points in development.

General Discussion

In the research reported above, we developed a modified (3-D) preferential looking paradigm which allows the study of very early word learning. Indeed, the results suggest not only that 12-month-olds can participate in the 3-D task, but also that we can begin to unpack theories about the processes children use in acquiring their very first words. While these results are only preliminary, they do have at least two implications for the controlled study of word learning.

First, it has proven very difficult to study early language acquisition. With a few notable exceptions, the vast majority of the word learning literature focuses on children three and four years of age. The introduction of the 3-D intermodal preferential looking paradigm joins methods like those used by Waxman and Markow (1995) or Woodward, Markman, and Fitzsimmons (1994), to provide a balanced and controlled procedure for use with infants as young as 12 months of age.

The second message from this program of research is that shifting focus downward suggests that seemingly incompatible approaches to the field might all be necessary to explain word learning. Thus, the social/pragmatic theorists need the constraints theorists. For without constraints, children would have difficulty figuring out exactly what is to be labeled once joint attention is established. (Is the speaker labeling the whole object, it's color, or the surface on which it rests?) Likewise, the constraints theorists need the social/pragmatics theorist. Without social/pragmatic input, children would have no idea which of the many unnamed whole objects around them the speaker intends to be the referent. That is, the social cues establish the where, and the constraints establish the what.

Why has the advantage of this polygamous union not been seen before? Partly because researchers prefer to work from a particular theoretical vantage point, and partly because the methods previously available prohibited careful study of very early word learning. Thus, researchers often began their studies after children had integrated many of the cues in the coalition available to them. In this paper, we have provided evidence of the value of one new method with which to peer into this early developmental period: the 3-D Intermodal Preferential Looking Paradigm. We believe that future study will ultimately demonstrate the contributions that each apparently contradictory theory makes to word learning over development. We believe that it is only in the coalition of cues (perceptual, social, temporal and linguistic), acting in concert, that children come to move from immature to mature lexical principles.

In conclusion, we used these studies of early word learning to highlight and develop a theory and a new method for examining word learning in children who are just beginning to acquire a vocabulary. It appears the 3-D paradigm works: children as young as 11 months are clearly looking in the direction of requested familiar stimuli. In addition, post-vocabulary spurt children are able to rapidly learn novel labels for novel objects in very few repetitions. Finally, preliminary results suggest that as infants develop they seem to shift their attention from perceptual cues to social cues, ultimately using social mentors as a guide to word learning.

The landscape of developmental psychology is changing. There are now a number of theories that focus on developmental behaviors as emergent phenomena -- in language, spatial knowledge, number theory, representation, etc. In such theories infants are thought to have the ability to attend to multiple cues in the input and to detect statistical co-occurrences (Saffran & Newport, 1996; Aslin, Woodward, LaMendola, & Bever, 1996; Morgan, Shi, & Allopena, 1996; Smith, 1995). In such theories, systems appear to organize and reorganize as children differentially weight certain inputs over others across development. Infants can not only detect these coalitions, but can use them to construct more complex behaviors. To investigate this new crop of theories, and to see how pieces of the coalition come together, we will need to look at behaviors earlier in development than we have been. The study described here is a first step in that direction. The 3-D intermodal preferential looking paradigm provides researchers with a powerful tool for investigating the emergent processes implicated in word learning.

References

Aslin, R., Woodward, J., LaMendola, & Bever, T. (1996). Models of word segmentation in fluent maternal speech to infants. In J. M. &. K. Demuth (Eds.), Signal to Syntax (pp. 117-135). Cambridge, MA: MIT Press.

Baldwin, D. A. (1991). Infants' contributions to the achievement of joint reference. Child Development, 62, 875-890.

Baldwin, D. A. (1993). Infant's ability to consult the speaker for clues to word reference. Journal of Child Language, 20, 377-394.

Baldwin, D. A. (in press). Early referential understanding: Infant's ability to recognize referential acts for what they are. Developmental Psychology.

Baldwin, D. A. & Moses, L. J. (in press). Early understanding of referential intent and attentional focus: Evidence from language and emotion.

Bloom, L. (1993). The Transition from Infancy to Language: Acquiring the Power of Expression. Cambridge, MA: Cambridge University Press.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan & G. A. Miller (Ed.), Linguistic Theory and Psychological Reality. Cambridge, MA: The MIT Press.

Carey, S. & Bartlett, E. (1978). Acquiring a single new word. Papers and reports on Child Language Development (Department of Linguistics, Stanford University), 15, 17-29.

Fagan, J. (1971). Infant recognition memory for a series of visual stimuli. Journal of Experimental Child Psychology, 11, 244-250.

Fagan, J., Singer, L., Montie, J. & Shepard, P. (1986). Selective screening device for the early detection of normal or delayed cognitive development in infants at risk for later mental retardation. Pediatrics, 78, 1021-1026.

Golinkoff, R. M. & Hirsh-Pasek, K. (1995). Reinterpreting children's sentence comprehension: Toward a new framework. In P. Fletcher and B. MacWhinney (Eds.), The Handbook of Child Language (pp. 430-461). London: Blackwell.

Golinkoff, R., Hirsh-Pasek, K., Cauley, K. & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. Journal of Child Language, 14, 23-46.

Golinkoff, R. M., Mervis, C. & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. Journal of Child Language, 21, 125-155.

Hirsh-Pasek, K. & Golinkoff, R. (1991). Language comprehension: A new look a some old themes. In N. Krasnegor, D. Rumbaugh, M. Studdert-Kennedy & R. Schiefelbusch (Eds.), Biological and behavioral aspects of language acquisition. Hillsdale, NJ: Erlbaum.

Hirsh-Pasek, K. & Golinkoff, R. (1993). Skeletal supports for grammatical learning: What infants bring to the language learning task. In C. Rovee-Collier & L. Lipsitt (Eds.), Advances in infancy research, (Vol. 8, pp. 299-338). Norwood, NJ: Ablex Publishing Corp.

Hirsh-Pasek, K. & Golinkoff, R. (1996a). The intermodal preferential looking paradigm reveals emergent language comprehension. In D. McDaniel, C. McKee, & H. Cairns (Eds.), Methods for assessing childrens' syntax. Cambridge, MA: MIT Press.

Hirsh-Pasek, K. & Golinkoff, R. (1996b). The origins of grammar. Cambridge, MA: MIT Press.

Hollich, Hirsh-Pasek, and Golinkoff, (in preparation) Children’s use of social cues in word learning.

Markman, E. M. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.), Concepts and Conceptual Development: Ecological and intellectual factors in categorization (pp. 255-287). Cambridge, MA: Cambridge University Press.

Markman, E. M. (1992). Constraints on word learning: Speculations about their nature, origin, and domain specificity. In M. R. Gunnar & M. P. Maratsos (Eds.), Minnesota Symposium on Child Psychology, 25 . Hillsdale, NJ: Erlbaum.

McDaniel, D., McKee, C. & Cairns, H. (Eds.). (1996). Methods for assessing childrens' syntax. Cambridge, MA: MIT Press.

Morgan, J., Shi, R. & Allopena, P. (1996). Perceptual bases of rudimentary grammatical categories: Towards a broader conceptualization of bootstrapping. In J. Morgan & K. Demuth (Eds.), Signal to Syntax (pp. 263-287). Cambridge, MA: MIT Press.

Smith, L. B. (1989). A model of perceptual classification in children and adults. Psychological Review, 96, 125-144.

Smith, L. B. (1995). Self-organizing processes in learning to learn words: Development is not induction. In C. A. Nelson (Ed.), Basic and Applied Perspectives on Learning, Cognition, and Development: The Minnesota Symposia on Child Psychology, (Vol. 28, pp. 1-32). Mahwah, NJ: Erlbaum.

Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. Developmental Psychology, 15, 626-636.

Tomasello, M. T. & Akhtar, N. (1995). Two-year olds use pragmatic cues to differentiate reference to objects and actions. Cognitive Development, 10, 201-224.

Tomasello, M. & Farrar, M. (1984). Cognitive bases of lexical development: Object permanence and relational words. Journal of Child Language, 11, 477-493.

Tomasello, M. & Kruger, A. C. (1992). Joint attention on actions: Acquiring words in ostensive and non-ostensive contexts. Journal of Child Language, 19, 313-333.

Tomasello, M. & Todd, J. (1983). Joint attention and lexical acquisition style. First Language, 4, 197-212.

Tucker, M., Hirsh-Pasek, K, Hollich, G. and Golinkoff, R. (in press) A change is afoot: Emergentist theories in language acquisition. In Anderson, P. (Ed.) Downward Causation.

Waxman, S.R., & Markow, D.B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. Cognitive Psychology, 29, 257-303.

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

Author Note

George J. Hollich & Kathy Hirsh-Pasek, Department of Psychology; Roberta Michnick Golinkoff, Department of Educational Studies.

This research was supported by NSF grant (#SBR9601306) to the second and third authors. The principle experiment reported herein was presented at the Society for Research in Child Development Conference, Washington, D.C., in April, 1997. Many former undergraduate, graduate, and post-graduate students have assisted over the years, and we would like to collectively thank them all. Particular thanks goes to Rebbecca Brand, Cindy Hankey, Camille Rocroi, and Judy Wiley.

Correspondence concerning this chapter should be addressed to George J. Hollich, Department of Psychology, Temple University, Philadelphia, Pennsylvania 19122. Electronic mail may be sent via Internet to ghollich@astro.ocis.temple.edu.

Figure Captions

Figure 1. The Intermodal Preferential Looking Paradigm. Children's looking times to the video monitors reflect preferences for displays which match the audio stimulus.

Figure 2. The Three Dimensional Intermodal Preferential Looking Paradigm. The presence of a live experimenter permits highly engaging and realistic labeling situations. Children's looking times to the objects on the display board reflect preferences for objects which match the auditory stimulus. Off-line coding is possible through video recording. The mirror allows simultaneous recording of the infant, the experimenter, and the stimuli.

Figure 3. The Modified Fagan Apparatus. The hinged 40 cm x 50 cm x 3 cm board allows for rapid hiding and revealing of the test objects.

Figure 4. The Phases of the Familiar Trials.

Figure 5. The Familiar Test Trials. Mean percentage of looking time to the requested object (the target) by age group. A score of .50 would indicate equal looking to both objects.

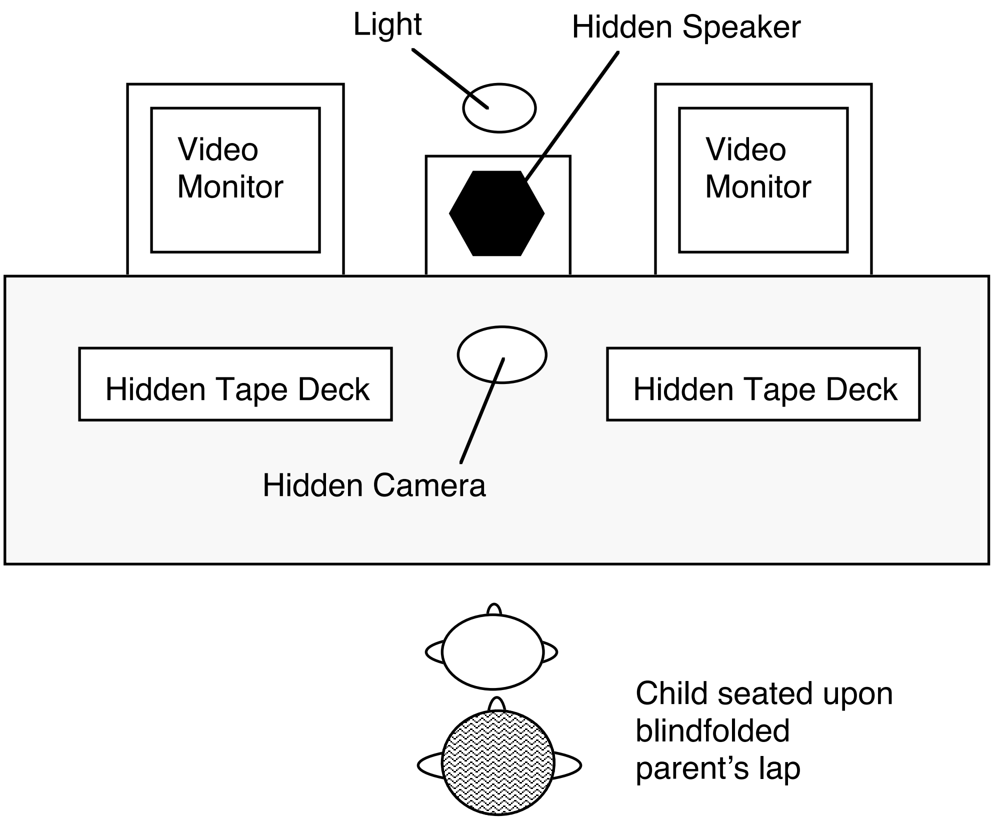
Figure 6. The Coalition Model for Word Learning. The child learner is surrounded by a coalition of cues, the weighting of which changes with time.

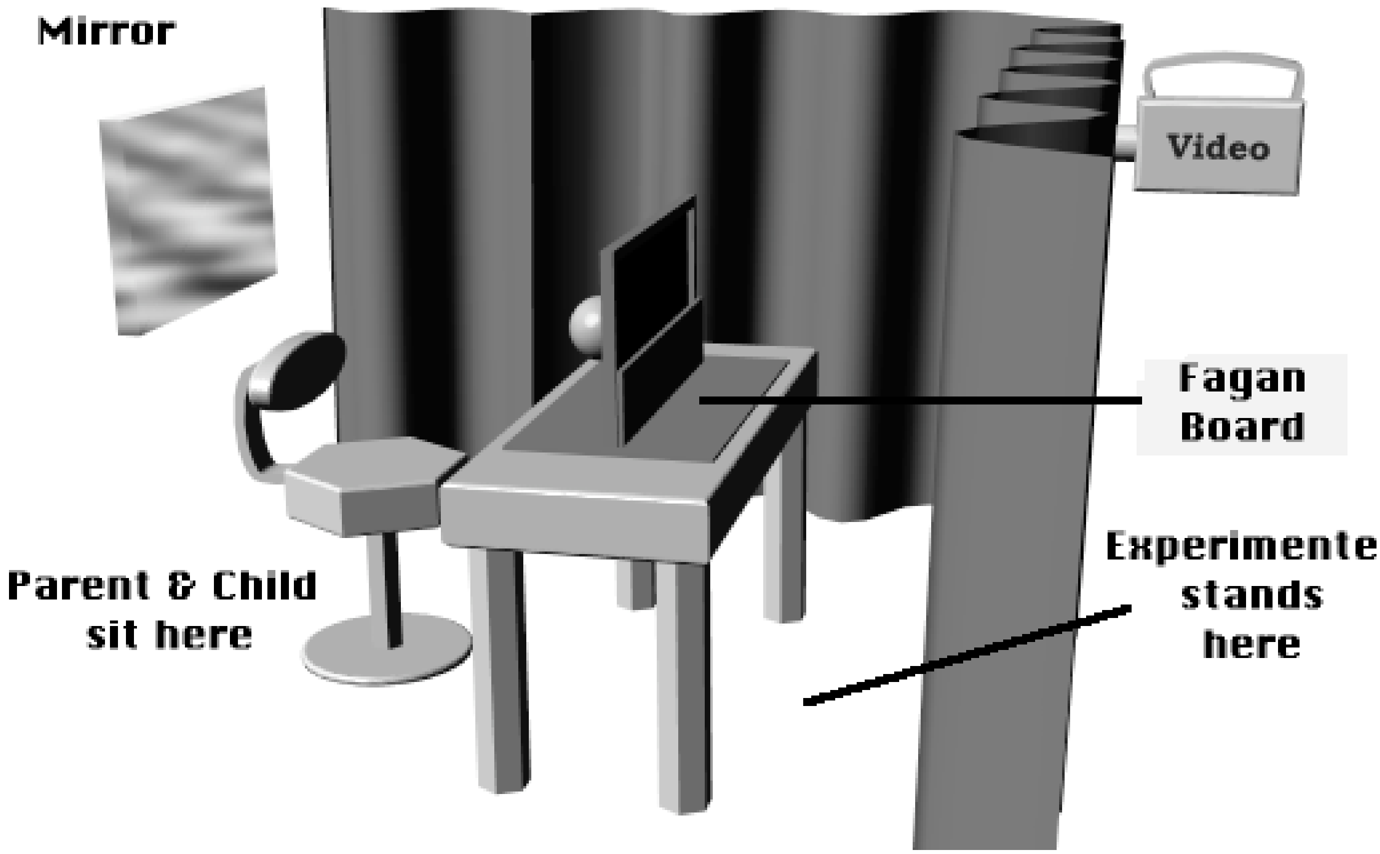
Figure 7. The Conditions of the Reference Study. A. The conflict condition: The boring toy is labeled while the experimenter attempts to draw the child's attention to the boring toy through the use of eye gaze. B. The coincidental condition: The interesting toy is labeled while the experimenter attempts to draw the child's attention to the boring toy through the use of eye gaze.

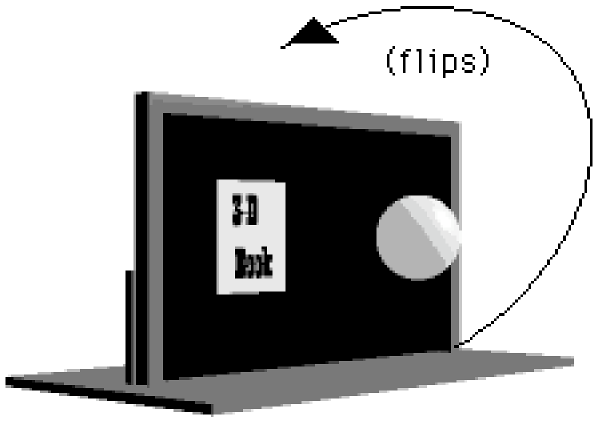
Figure 8. The Phases of the Novel Trials.

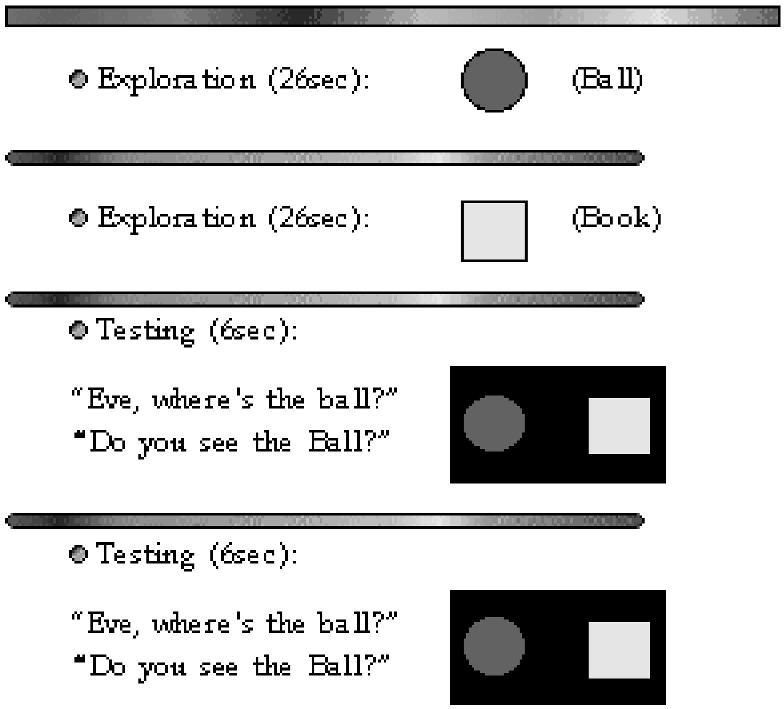
Figure 9. Salience and Training Trials. Mean percentage of looking time to the interesting object by age, salience score, and condition. A score of .50 would indicate equal looking to both objects.

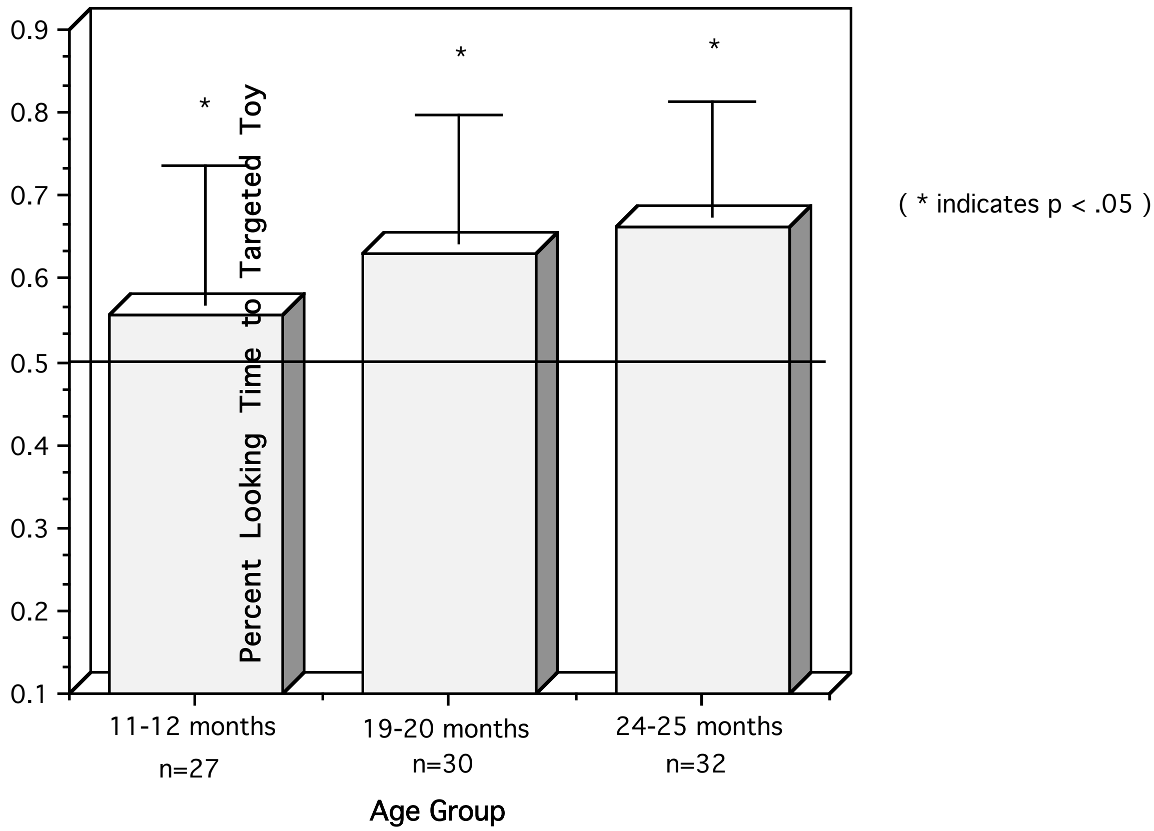
Figure 10. Salience and Test Trials. Mean percentage of looking times to the interesting object by age, salience score, and condition. A score of .50 would indicate equal looking to both objects.

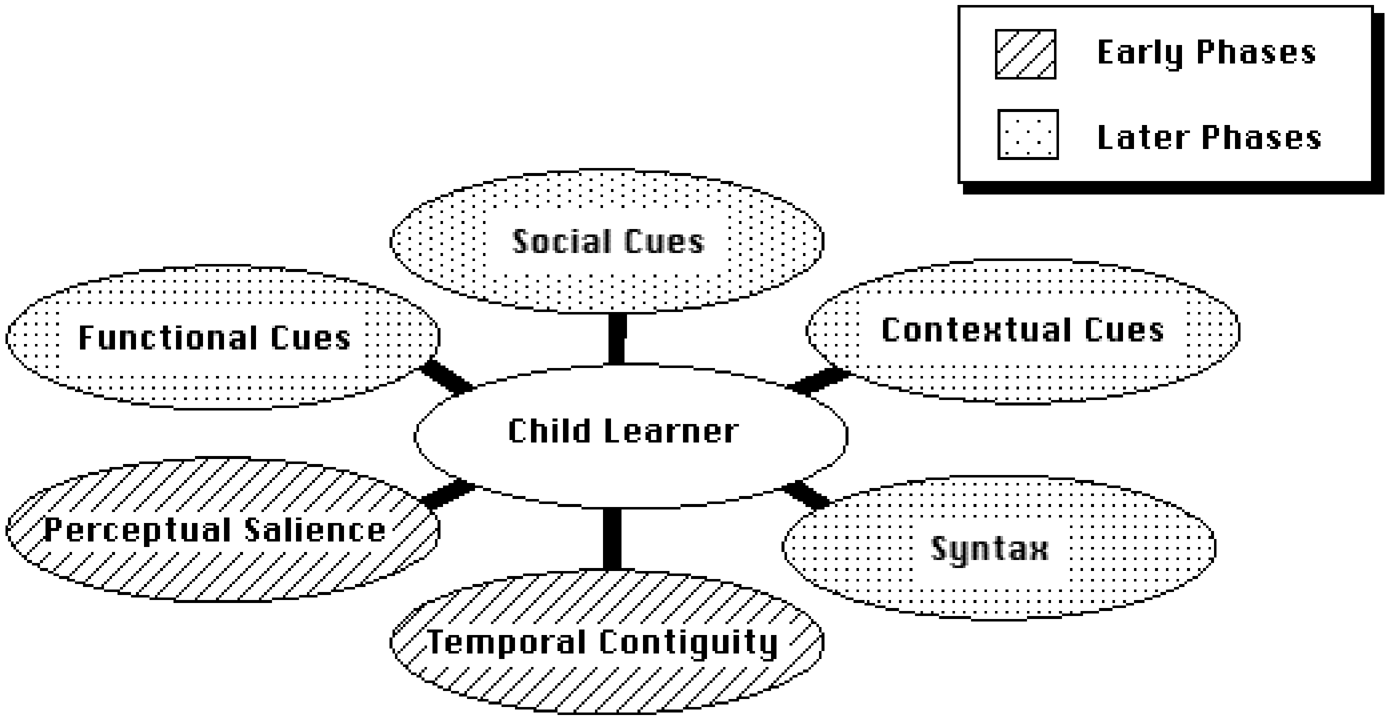

Light
Hidden Speaker
Video
Monitor
Video
Monitor
Hidden Tape Deck
Hidden Tape Deck
Hidden Camera
Child seated upon
blindfolded
parent's lap

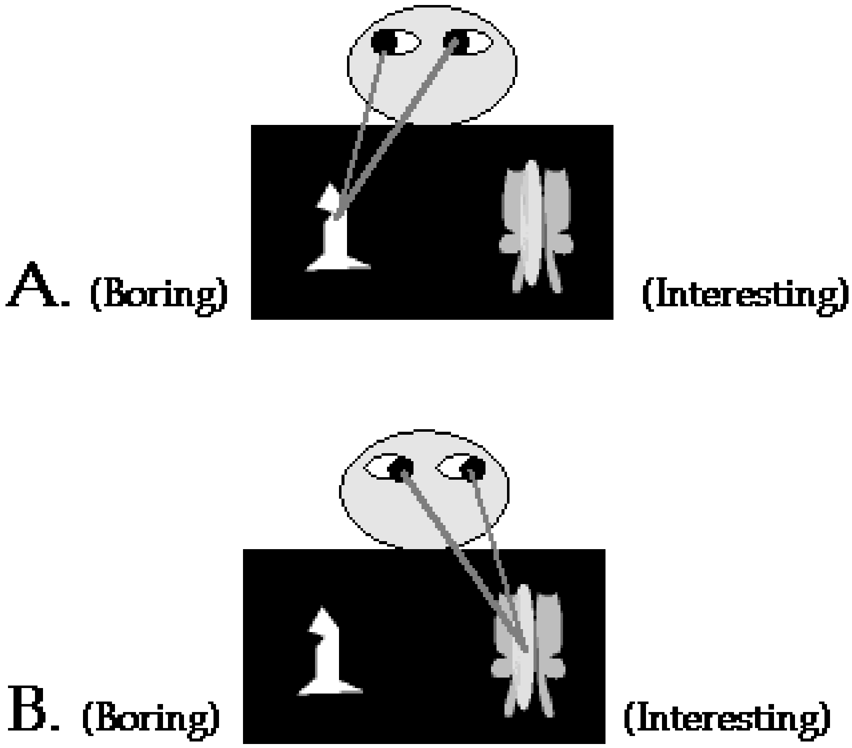

Mirror
Video
Fagan Board
Parent & Child sit here
Experimenter stands here


(flips)
3-D
Book

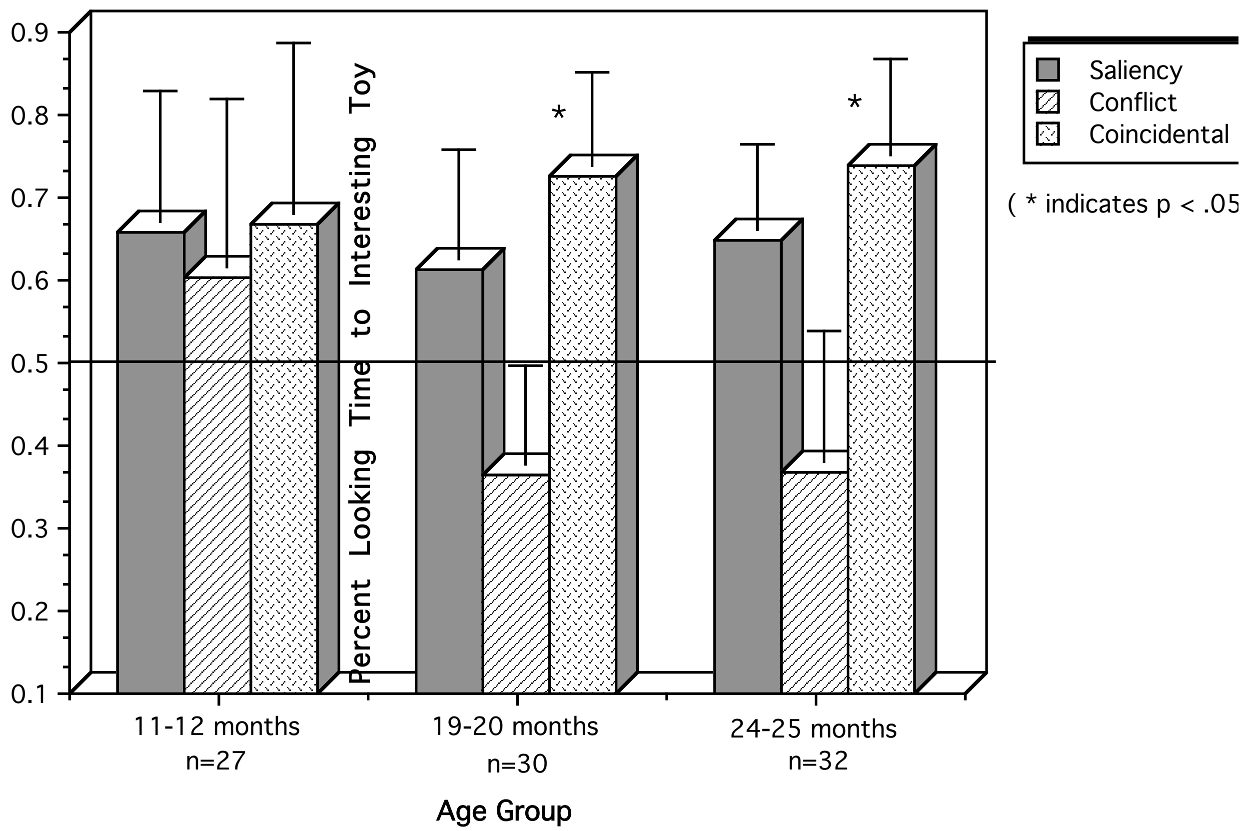

Exploration (26sec):
(Ball)
Exploration (26sec):
(Book)
Testing (6sec):
"Eve, where's the ball?"
"Do you see the Ball?"
Testing (6sec):
"Eve, where's the ball?"
"Do you see the Ball?"

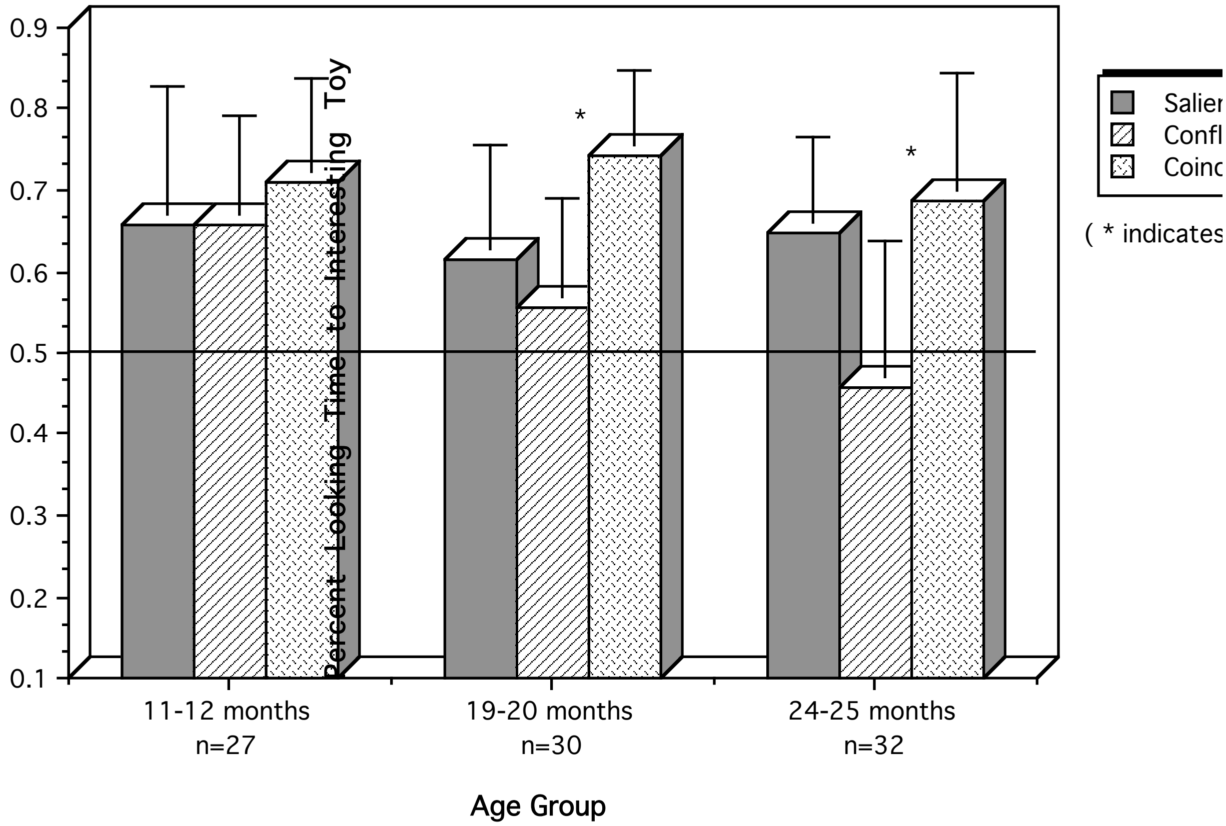

0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Percent Looking Time to Targeted Toy
*
*
*
11-12 months
n=27
19-20 months
n=30
24-25 months
n=32
Age Group
(* indicates p < .05)


Early Phases
Later Phases
Social Cues
Functional Cues
Contextual Cues
Child Learner
Perceptual Salience
Syntax
Temporal Contiguity


A. (Boring)
(Interesting)
B. (Boring)
(Interesting)


Exploration (26sec):
(Boring)
Exploration (26sec):
(Interesting)
Salience (6sec):
"Look at that!"
"What do you see?"
Training (≈16sec):
"Eve, look at the danu!"
"A danu! See the danu!"
Testing (6sec):
"Eve, where's the danu?"
"Look at the danu."
Testing (6sec):
"Eve, where's the danu?"
"Look at the danu."


0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Percent Looking Time to Interesting Toy
Saliency
Conflict
Coincidental
(* indicates p < .05
*
*
11-12 months
n=27
19-20 months
n=30
24-25 months
n=32
Age Group


0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Percent Looking Time to Interesting Toy
*
*
11-12 months
n=27
19-20 months
n=30
24-25 months
n=32
Age Group
Salier
Confl
Coinc
(* indicates