

Referential Understanding of Videos in Chimpanzees (*Pan troglodytes*), Orangutans (*Pongo pygmaeus*), and Children (*Homo sapiens*)

Sarah R. Poss and Philippe Rochat
Emory University

Performance on identical search tasks based on cues directly perceived or indirectly perceived through video were compared among a group of 4 adult chimpanzees (*Pan troglodytes*), a group of 2 adult orangutans (*Pongo pygmaeus*), and a group of 36 children (between 2 and 3 years of age). Children comprehended directly perceived cues but had difficulty with video cues. In contrast, chimpanzees and 1 orangutan were successful in using video to guide their search for a hidden object. Two follow-up studies with 3-year-old children demonstrated the importance of more distinct perceptual and verbal cues in aiding children's understanding of video as referring to real-world events.

Human infants grow up in an environment that is filled with referential artifacts such as television, pictures, and books. From a very young age and across most cultures, children are nurtured for their propensity to engage in symbolic and pretense play. Throughout this period, they are provided with toys and other replicas to reenact real-world objects and events (Rochat, 2001; Striano, Tomasello, & Rochat, 2001; Tomasello, Striano, & Rochat, 1999).

The question of interest here is when do children begin to understand and use the referential nature of artifacts to guide their action in the real-world environment? Furthermore, we ask whether nonhuman primates, who are not typically surrounded and nurtured with symbolic artifacts, might also show an ability to comprehend symbolic artifacts such as video images as referring to real-world events. Our question pertains to the origin of an understanding of virtual images as representing real things and events in the world. Here, we compare young children and two great ape species (chimpanzees and orangutans) in an identical search task guided by virtual cues displayed on a television screen. By adopting a comparative perspective, we hope to provide new information on what might determine the development of referential understanding in humans as compared with nonhuman primates.

Developmental psychologists as well as primatologists have investigated the referential understanding of artifacts (e.g., replicas, scale models, television, or photographs) in children and apes as expressions of symbolic competence (DeLoache, 1991, 2000; Kuhlmeier & Boysen, 2001, 2002; Kuhlmeier, Boysen, & Mukobi,

1999; Namy & Waxman, 1998; Schmitt & Anderson, 2002; Tomasello, Call, & Gluckman, 1997; Troseth & DeLoache, 1998). Developmental research indicates that by 2.5 years of age, children begin to manifest a referential understanding of pictures, including videos (Troseth & DeLoache, 1998). In Troseth and DeLoache's (1998) research, children saw an object being hidden in an adjacent room via a video. They were then asked to search for the object in the room. This paradigm was used to test children's ability to navigate and guide their search in real space on the basis of video cues (see also Schmitt & Anderson, 2002).

If children by 2.5 years of age are capable of using video cues as referential to real-world events, this capacity might not be specific to humans. Many studies have been done illustrating that nonhuman primates are capable of visible displacement tasks. In particular, extensive tests have been conducted with apes showing their complex skills at finding objects in spatial cognition tasks (Hallock & Worobey, 1984; Mathieu & Bergeron, 1981; Natale, 1989; Redshaw, 1978). However, the extent to which nonhumans are capable of using items in their environment as a representational tool to guide them in these searches and the degree to which this ability is similar to that in human children is the current issue. Symbolic functioning in nonhuman primates has been addressed through several studies involving great apes. These nonhuman primates have shown a capacity to use symbols such as American Sign Language and visual graphics to communicate with their caregivers (Fouts, 1998; Gardner, Gardner, & Van Cantfort, 1989; Premack, 1986; Savage-Rumbaugh, McDonald, Sevcik, Hopkins, & Rupert, 1986). Although some of this research and its interpretation raise much controversy (Seidenberg & Petitto, 1987), it is difficult to ignore the mounting evidence that these studies show in favor of some symbolic comprehension by nonhuman primates.

In a study done by Menzel, Premack, and Woodruff (1978), infant chimpanzees (about 3 years of age) were tested for their ability to locate caregivers in an enclosure after watching them hide on a video screen. Infant chimpanzees watched on a small, black-and-white television monitor as their caretaker walked into an outdoor enclosure. The caretaker then proceeded to a particular location in the enclosure and hid there. At this point, the infant chimpanzees were released into the enclosure to search for the caretaker. The view from the monitor was such as to simulate the

Sarah R. Poss and Philippe Rochat, Department of Psychology, Yerkes Regional Primate Research Center, and Living Links Center, Emory University.

This investigation was supported in part by an Emory University Research Committee Grant (99–00). We thank Chuck Brady, Matt Thompson, and the Memphis Zoo for their valuable cooperation and for allowing us to test the orangutans. We also thank William Hopkins for his assistance and the numerous undergraduate research assistants who helped with data collection and coding. The Yerkes Regional Primate Research Center is fully accredited by the American Association of Laboratory Animal Care.

Correspondence concerning this article should be addressed to Sarah R. Poss, Department of Psychology, Emory University, 532 North Kilgo Circle, Atlanta, Georgia 30322. E-mail: sposs@emory.edu

point-of-view of the chimpanzee if looking through a window into the enclosure. Results indicate that chimpanzees were capable of such a search using the live video cues to represent the actual enclosure. This study illustrates an ability of chimpanzees to transpose information from a video support to their real-world referents. Vauclair (1990) offered similar observations in adult Guinea baboons (*Papio papio*).

In studies done by Kuhlmeier and colleagues (Kuhlmeier et al., 1999; Kuhlmeier & Boysen, 2001, 2002), attempts were made to replicate DeLoache's findings using scale models with a group of chimpanzees (*Pan troglodytes*). Kuhlmeier and colleagues tested the abilities of chimpanzees to use scale models as symbolic of a real-world environment. Results indicate that chimpanzees are sensitive to object and positional cues, much like young children, when mapping the model to its referent. Some differences between the two species appear to be a more rigid adherence to specific search patterns on the part of the chimpanzees. Although these studies illustrate that chimpanzees are capable of using information from a referential, three-dimensional object such as a scale model, there are still differences and similarities in the ways in which children and chimpanzees use cues to make such a connection. Kuhlmeier and Boysen (2002) pointed out that chimps and 3-year-old children show a sensitivity to both object and spatial correspondences. However, as evidenced by the strong performance of 2 chimpanzees in a condition in which cues were completely absent (a condition in which 3-year-old children have difficulty), chimpanzees might more readily recognize spatial correspondences than 3-year-old children (see Kuhlmeier & Boysen, 2002, pp. 62–63).

Questions remain regarding the competence for representational understanding underlying the performance of apes or children alike. Does the behavior observed necessarily rely on referential (symbolic) understanding, or could it be supported by literal (direct) processing of perceptual information or by stimulus enhancement? In addition, how does this competence compare across closely related primate species? The aim of the present research is to address these questions by comparing the performance of human children, chimpanzees, and orangutans on an identical search task that seems referential.

Existing studies comparing human and nonhuman cognitive abilities typically proceed by trying to replicate performance in animals in the context of tasks that are first and foremost designed for human children (e.g., Call & Rochat, 1996, 1997; Kuhlmeier et al., 1999; see Tomasello & Call, 1997). This approach tends to put children at an advantage by taking for granted the complex social and linguistic scaffolding these tasks entail—in particular, the necessity of verbal instructions. As a case in point, many developmental research studies on humans show that young children's symbolic performance in nonlinguistic domains is highly dependent on verbal instructions (Callaghan, 2000; DeLoache, 1991) and the verbal parsing of the task (Gentner & Rattermann, 1991).

To control for this caveat, we proceeded in comparing referential capabilities in human and nonhumans by designing a task that is suitable for nonlinguistic primate species, namely, a task that minimizes as much as possible explicit verbal instructions. Accordingly, we reverse what is usually done in comparative cognition research. We designed a task that was first adapted to nonhuman primates, then used the exact same task with human children for comparison. The procedure and design attempt to

control for local (or stimulus) enhancement for the nonhuman primates, as well as verbal instructions for the human children. In general, all participants are tested in a task in which an object (food or toy) is hidden inside one of two objects, removed from sight, and then presented again for search. In one condition, the reward is hidden directly in front of the participant to establish a baseline understanding of the hiding task. In a video condition, participants view the hiding event indirectly on a large TV screen and are then presented with the two objects for search in real space. In addition, to control for possible local enhancement, we set up the experiment so that participants either do or do not see the actual hiding of the object. In other words, in one event, the experimenter places the reward in the target object in front of the participant. In the other event, the experimenter simply shakes the target object or hiding location in front of the participant but does not demonstrate the actual hiding. We compare search performance in each of these hiding conditions. The following two working hypotheses guided the research.

1. On the basis of previous developmental research (i.e., Schmitt & Anderson, 2002; Troseth & DeLoache, 1998), we expected that by 2.5 years (30 months) of age, children would be successful in using video to assist them in searching for the reward. Considering existing primate literature (Menzel, 1971; Menzel et al., 1978; Vauclair, 1990), we also expected that, overall, the performance of nonhuman primates would be comparable to 2.5-year-old children in the video condition compared with the direct condition.

2. Despite their comparable performance, we expected that nonhuman participants would tend to rely more on stimulus enhancement cues. Stimulus and local enhancement refer to animals' propensity to prefer the locations occupied by other animals or objects previously manipulated by them (Thorpe, 1956; Tomasello & Call, 1997). In the context of our experiment, if participants (nonhuman or human) are relying on stimulus enhancement cues, we would expect to see no difference in performance between conditions in which they witnessed the actual baiting versus the conditions in which they did not. This rationale is based on the fact that in both conditions (visible hiding and invisible hiding; see method below), the target object was manipulated more by the experimenter than the other, whether or not the participant saw the actual hiding of the reward in the manipulation. In other words, the second hypothesis was that for nonhuman participants only, the performance would not differ significantly between the latter conditions. The rationale for these hypotheses was that nonhuman participants would be sophisticated perceptual analysts (see Call & Rochat, 1997) capable of using subtle perceptual cues such as stimulus enhancement to guide their search for the reward. Thus, the simple act of manipulating one location more than another should be enough to guide their search strategies.

General Method

All participants were tested in four search-task conditions: (a) direct, visible hiding; (b) direct, invisible hiding; (c) video, visible hiding; and (d) video, invisible hiding. In the direct conditions, the participant faced the experimenter at all times during the task. In the video conditions, the participant faced the experimenter during the search phase of the task but viewed the hiding event on a video monitor. In the visible-hiding conditions, the participant witnessed the actual placement of the reward (food or small doll) in one of the two hiding objects. In the invisible-hiding conditions, the participant witnessed the experimenter shaking the object in

which the reward was eventually hidden. In other words, both objects were held up in front of the participant, but the object in which the reward was to be hidden was shaken. In all conditions, the experimenter manipulated both hiding objects about equally, and both were always present at the same time. This procedure was the same for all experiments.

Experiment 1

Method

Participants. Four subadult chimpanzees (*Pan troglodytes*), 2 adult orangutans (*Pongo pygmaeus*), and 36 human children (*Homo sapiens*) were tested. The group of chimpanzees consisted of 3 females and 1 male, ranging in age from 6–12 years ($M = 8.25$ years, $SD = 2.63$). These participants were housed in social groups at the Yerkes Regional Primate Research Center. One female (Wilma) was mother reared, and the other participants (Faye, Julie, and Patrick) were nursery reared. Inclusion in this study required no previous training of the participants. All chimpanzees had previous experience as participants in other cognitive testing; however, none of them had been exposed to video prior to this task. The procedure was new to all of these animals, and they did not require any preliminary training. Testing was conducted in the participants' indoor home cages. For all participants, testing was conducted in the presence of only one other cage mate. This cage mate was distracted with food items so that interference would not be a factor during testing. Julie and Patrick were the only 2 participants in the study who were housed together. In their case, an additional experimenter sat outside the cage during testing to feed and distract the other participant to ensure that their performances would be independent of one another. There was no apparent interaction between animals. The animals were fed twice daily on a diet of monkey chow, fruit, and vegetables. These animals were never food deprived, and water was available to them ad libitum. In addition to the participants included in the study, we attempted to test another adult female chimpanzee (Lee) in the first condition but terminated testing because of her inability to display unambiguous choice.

Two adult orangutans (*Pongo pygmaeus*) housed at the Memphis Zoo were also tested (Puti & Tombak; see further description in Call & Rochat, 1997). Both had been nursery reared. The procedure was new to both animals, and they did not require preliminary training. They lived in an indoor and outdoor area and were tested in the former. They were fed twice daily on a diet of monkey chow, fruit, and vegetables; they were never food deprived; and water was available to them ad libitum. In addition to the 2 participants included in the study, we attempted to test another adult female orangutan (Chickie) in the direct condition but terminated testing because of her overwhelming right-side preference.

Children were 24, 30, and 36 months of age, with 12 in each age group and a total of 18 boys and 18 girls. They were recruited from a participant pool consisting of children from the Atlanta area. Parents were contacted by phone and invited to participate in the study. Participants were mostly White and middle class. They received a small gift for their participation. Parents were allowed to remain present during all phases of testing. The children were tested at the Child Study Center at Emory University. In addition, some of the children were recruited and tested at Atlanta-area day care centers (Decatur First United Methodist Children's Morning Out Program and Weiucca Baptist Church Day Care Center). Signed consent forms were obtained from parents of all children. We tested 14 additional children (nine 24-month-olds, four 30-month-olds, and one 36-month-old). These children were excluded because of fatigue and fussiness.

Apparatus. The apparatus was the same for all groups of participants. For hiding objects, we used polyvinyl chloride (PVC) tubes (2.5-in. [6.35-cm] diameter) of varying colors (red, yellow, and blue) and with varying shape markers (stripes, dots, or swirls). The tubes had a piece of connector PVC piping at the end such that the tube could be slipped through the bars of the cage for the animal participants, without being completely pulled inside. These tubes and this basic technique were previously used to test

hand preference in chimpanzees (Hopkins et al., 2001). Objects hidden were food items (pieces of banana) for nonhumans and a small stuffed bear (approximately 2.5 in. [6.35 cm] in height) for the children. In addition to these objects as reward, all participants received praise and verbal encouragement following correct choices.

Nonhumans performed the search task in their home cage. Children were tested in a quiet room and sat on a small chair at a low table. In direct conditions, all participants viewed the experimenter hide an object in one of two tubes, in front of them. In the video conditions, all participants saw the experimenter hide the object in one of two tubes on a video monitor (19 in. [48.3 cm]) placed approximately 3–4 feet (0.9–1.2 m) away. While hiding the object, the experimenter stood behind an opaque white curtain hung behind the monitor and was filmed online by a Panasonic AG187 video camcorder.

Procedure. Aside from the kind of reward (food or stuffed animal) and the testing location (home cage vs. testing room with furniture), the procedure was identical for all participants.

Children were tested in a single session with 6 trials per condition for a total of 24 trials. Nonhumans were tested in two sessions over 2 separate days with 24 trials per condition and two conditions per day, for a total of 96 trials. On the basis of pilot observations, children were tested with fewer trials to alleviate fatigue and boredom effects. The greater number of trials per condition for nonhumans compensated for the fewer available number of participants. Note that 1 chimpanzee (Faye) was tested on only 20 trials per condition because of her commitment to other research at Yerkes.

All participants were tested first in the direct, visible-hiding condition, followed by the direct, invisible-hiding condition. The rationale for starting with the direct, visible-hiding condition was to introduce to participants the nature of the search task. The order of the video, visible-hiding versus the video, invisible-hiding conditions was counterbalanced among participants of each group. The left–right presentation of hiding location was randomized.

Between direct and video conditions, all participants were presented with a video orientation phase during which they had a 60-s opportunity to monitor themselves online on the video screen. This was done in an attempt to familiarize the participants with the apparatus.

Because nonhumans were tested in two sessions over 2 days, at the beginning of the second session, participants were retested in one of the direct conditions as a warm-up reminder of the search task. This warm-up ended when the animal performed successfully on four consecutive search trials.

In this first direct condition (direct, visible hiding), participants directly observed the experimenter place the reward in one of the two tubes. The hiding took approximately 2 s. Both tubes were equally manipulated (i.e., touched and handled) by the experimenter in view of the participant prior to the hiding. This was done to minimize the chance that participants would only pay attention to the tube manipulated by the experimenter. The reward was always hidden in front of the participant in front of the experimenter's body. The experimenter then placed the tubes behind her back for approximately 2 s. Both tubes were then presented in a predetermined random left–right position in front of the participant about 15 in. (38.1 cm) apart and 10 in. (25.4 cm) away from the participant, within reaching distance in a fronto-parallel presentation. The experimenter stared straight ahead at the individual to ensure that no visual cues were given to the participant. For nonhumans, tubes were held with the opening up, so that participants could not look down inside to see the reward. In addition, to control for olfactory cues, we made sure that both tubes were previously rubbed with banana. For children, tubes were placed on the table with the opening facing away from the child to avoid direct perception of the stuffed animal.

In the second direct condition (direct, invisible hiding), the procedure was the same, except that the hiding occurred behind the experimenter's back prior to test presentation and following the gesturing. The experimenter held up both objects and shook one tube for approximately 2 s. The experimenter then placed both objects behind her back and hid the reward

inside of the object that she had shaken. The hiding took approximately 2 s. In this way, the participant did not actually witness the experimenter place the reward inside of the object.

In the video conditions, the experimenter hid behind a curtain, appearing live on the video screen placed 3 feet (0.91 m) away from the participant. The video screen was present during the direct conditions but was covered by a white sheet. When video conditions began, the white sheet was removed. The experimenter would walk behind a white curtain placed just behind the television and appear on the screen. The experimenter appeared on the screen from the waist up and performed as in the other direct, visible-hiding or invisible-hiding conditions. The experimenter reappeared from behind the curtain for the testing (choice) phase as in the direct conditions. In all conditions, verbal instructions were minimal for both children and nonhuman participants. The experimenter would ask the participant to watch what she was doing but gave no explicit instructions about the task. The experimenter would say, during the shaking or hiding portion, "Do you see this one?" Then, during the choice portion of the task, the experimenter would say, "Okay, which one do you think it is in?"

Coding and dependent measure. Choices were recorded on the basis of the first location the participant either pointed to or the first tube that was manually touched. Pointing was operationally defined as a manual orientation toward one of the two tubes with finger or whole-hand extension (see Hopkins et al., 2001). Touching entailed physical contact of one hand with a particular tube. In cases of bimanual engagement in which both hands oriented and eventually touched both tubes at the same time, the experimenter removed the tubes and presented them again.

For nonhumans, only correct choices resulted in a food reward in that the participant was given access to the hidden banana piece. In the case of an incorrect choice, participants were shown where the banana was hidden. Children were given the tube for toy removal when accurate. When a child pointed to or touched the wrong tube, the experimenter encouraged the child to look into the other, correct tube, which was given for toy removal. Although this method may have resulted in motivation differences for children, on the basis of pilot observations, this reward system was necessary to sustain children's engagement in the game (see children group's attrition rate). Thus, children were reinforced regardless of correct or incorrect choices.

For coding, two independent coders viewed videotapes of the respondents and recorded for each condition whether the participant displayed a correct or incorrect search response. In addition, the experimenter recorded the correct or incorrect response for each condition. Note that, because of the procedure, both observers knew where the reward was hidden. Percentage agreement between coders on correct and incorrect responses of all tested participants across all trials in each condition was above 99%. In addition, over 99% of the experimenter's online coding of correct and incorrect responses was in agreement with her own recoding of the video recording (high intraobserver reliability).

Results

Because of the relatively small number of participants, the data for nonhuman participants were analyzed using nonparametric statistics. The sign test was used to establish whether group performance differed significantly from chance (50%) for each condition. In addition, we conducted a Friedman analysis of variance (ANOVA) for repeated measures to assess differences in performance medians for within-subjects variables.

Chimpanzees. A Friedman test was conducted to determine differences in medians among the four experimental conditions: direct, visible hiding ($Mdn = .77$); direct, invisible hiding ($Mdn = .67$); video, visible hiding ($Mdn = .84$); and video, invisible hiding ($Mdn = .63$). The test was significant, $\chi^2(3, N = 4) = 9.15$, $p = .027$, with a Kendall coefficient of concordance of .763, indicating

a strong difference among the four conditions. Follow-up, pairwise comparisons were conducted to determine differences between medians using Wilcoxon tests. The median score for visible-hiding conditions was significantly higher than the median score for invisible-hiding conditions ($p = .012$), but the median score for video conditions did not differ significantly from the median score for direct conditions ($p = .944$). In addition, there were no significant differences among the medians for video, visible-hiding and direct, visible-hiding conditions ($p = .194$), nor among video, invisible-hiding and direct, invisible-hiding conditions ($p = .465$).

As shown in Figure 1, sign tests indicate that chimpanzees performed significantly above chance (50%) in all four experimental conditions (direct, visible hiding, $p < .005$; direct, invisible hiding, $p < .005$; video, visible hiding, $p < 1.6 \times 10^{-9}$; and video, invisible hiding, $p < .028$). Figure 1 illustrates overall mean percentage correct scores for chimpanzees in each condition. Table 1 presents each of the 4 participants' individual performance as a function of condition.

Orangutans. A Friedman test was conducted to determine differences in medians among the four experimental conditions: direct, visible hiding ($Mdn = .79$); direct, invisible hiding ($Mdn = .52$); video, visible hiding ($Mdn = .65$); and video, invisible hiding ($Mdn = .51$). The test was not significant, $\chi^2(3, N = 2) = 2.37$, $p = .50$, suggesting that there were no significant differences among the medians for the four conditions.

Sign tests revealed that orangutans performed significantly above chance as a group on only the direct, visible hiding condition ($p \leq .001$), as shown in Figure 2. Performance for this group was not significantly above chance for any other condition. Looking at individual performance, however, yields a significant result for Puti on the video, visible-hiding condition ($p \leq .007$). Table 2 presents the participants' individual performance.

Children. We analyzed the mean percentage of correct responses for children, using a repeated-measures, two-way ANOVA with hiding (visible vs. invisible) and video (video vs. direct) as within-subjects variables and with age being the between-subjects variable (24 months, 30 months, and 36 months). The ANOVA revealed only a main effect of video on children's performance, $F(1, 29) = 23.61$, $p < .001$ (video: $M = .49$, $SD = .03$; direct: $M = .66$, $SD = .04$). This indicates that performance for children in the direct conditions was significantly different than performance in video conditions. The analyses did not reveal a

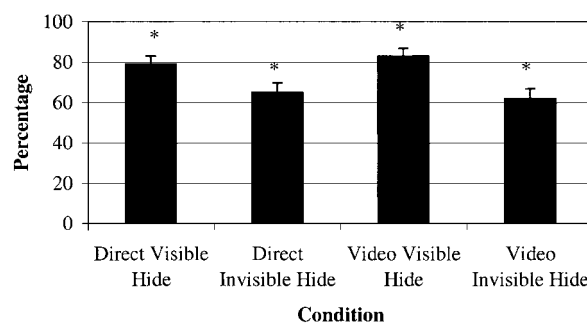


Figure 1. Mean percentage of correct responses for chimpanzees in Experiment 1 as a function of condition. Error bars reflect standard error as a function of mean percent correct. * $p < .05$.

Table 1
Chimpanzees: Number of Trials Correct Out of 24 Trials in Experiment 1

Chimpanzee	Condition			
	Direct, visible hiding	Direct, invisible hiding	Video, visible hiding	Video, invisible hiding
Wilma	16	15	18	16
Patrick	17	12	19	15
Julie	20	19	21	15
Faye ^a	19	14	18	11

^a Chimpanzee was tested in only 20 trials in each condition.

main effect of hiding nor any interaction among all possible combinations of age, hiding, or video conditions. Because no effect of age was found, analyses on the various age groups were not conducted.

As shown in Figure 3, sign tests indicate that children performed significantly above chance in only the direct, visible-hiding and direct, invisible-hiding conditions ($p \leq 2.75 \times 10^{-9}$ and $p \leq 6.5 \times 10^{-6}$, respectively).

Discussion

We expected that children by at least 30 months of age as well as chimpanzees and orangutans would perform equally well in the direct and video conditions. The results show that chimpanzees (and 1 orangutan) are the only group successful in both direct and video conditions in finding hidden objects at rates significantly higher than chance. It appears that children showed a significant deterioration of their search performance in the video conditions compared with the direct conditions. Chimpanzees tended to perform equally well in the video and direct conditions for visible-hiding conditions. In other words, compared with the group of human children, chimpanzees showed more versatility in using video cues to guide their search. One of the orangutans also seemed capable of using the video cues to guide her search. The question still remains as to what may account for this versatility.

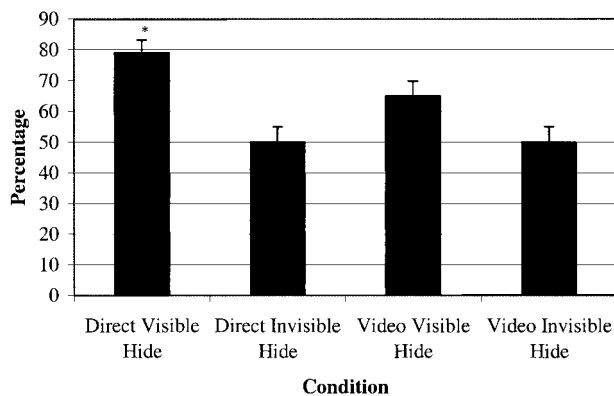


Figure 2. Mean percentage of correct responses for orangutans in Experiment 1 as a function of condition. Error bars reflect standard error as a function of mean percent correct. * $p < .05$.

Table 2
Orangutans: Number of Trials Correct Out of 24 Trials in Experiment 1

Orangutan	Condition			
	Direct, visible hiding	Direct, invisible hiding	Video, visible hiding	Video, invisible hiding
Tombak	21 ^a	12	12	14
Puti	17	13	19	9

^a Orangutan was tested in 27 trials in this condition.

We see two main possibilities. One is that, compared with children, great apes might be literal in their perception of the task, extracting and processing surface cues regarding the hiding place. The color of the tube and object manipulation by the experimenter would be treated at face value, whether perceived directly or on the TV screen. Presumably, this literal perception and processing of the hiding cues would avoid the problem of mapping the reality of the TV onto the real space in which the actual search takes place. Children, in contrast, would be less apt at such literal TV viewing. Research shows that by 2 years of age, children are facilitated in their search of an object they see by means of a TV being hidden in an adjacent room if they are made to believe that they are actually seeing the hiding event directly through a window (Trosseth & DeLoache, 1998). In other words, children were told that they were watching the hiding event through a window when, in fact, they were actually watching the event on a television screen. This method appeared to facilitate better performance in young children. One explanation for this increased performance could be that this method allows children to view what they are seeing on the television screen more literally. In addition, Schmitt and Anderson (2002) clearly demonstrate that up to 3 years of age, toddlers have difficulty using a video image to guide their search of an object in a real setting. Our results with children corroborate these observations.

Another possibility for the lesser performance of children in the video condition compared with the direct condition might be because for comparability across species, we designed an experimental situation that minimized verbal instructions and perceptual pointers, which may have put children at a disadvantage. In other words, the lesser performance of children could be due to what

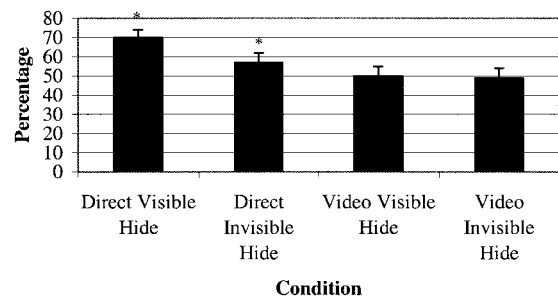


Figure 3. Mean percentage of correct responses for children in Experiment 1 as a function of condition. Error bars reflect standard error as a function of mean percent correct. * $p < .05$.

might amount for them to be the compounding effect of impoverished perceptual cues, the absence of verbal instructions, and less motivation to emit the correct response. The next two experiments consider specifically some of these possibilities by retesting groups of 3-year-old children in the same task and within the same design but providing them with enhanced perceptual and verbal cues.

Although apes may be literal in their perception of hiding, in general, it does not seem to be the case that they tended to rely more on stimulus enhancement cues to guide their search. Following our rationale, this would be the case if, compared with children, the apes demonstrated less differential performance in the visible-hiding compared with the invisible-hiding conditions. If stimulus enhancement were the strategy used by nonhumans, we would expect to see equivalent performance in both visible and invisible-hiding conditions because in each more attention is drawn to one hiding location than the other (by hiding a food reward inside of one for the visible-hiding condition and by shaking one location in the invisible-hiding location). In fact, as a group, chimpanzees performed significantly better in the visible-hiding compared with the invisible-hiding condition. This is also supported by the individual data for chimpanzees (see Table 1). From both the group and individual data, it is clear that the invisible-hiding conditions proved to be more difficult for chimpanzees than the visible-hiding conditions. Therefore, it appears that the chimpanzees based their search on more than mere stimulus enhancement cues.

In addition, the difference in performance for apes between visible and invisible-hiding conditions may be interpreted as a lack of referential understanding on the part of the nonhuman participants. It may be argued that the invisible-hiding conditions are more referential because they require the understanding that the shaking gesture by the experimenter is meant to indicate that something is to be hidden there. Perhaps this is too abstract for the apes to appreciate and hence results in the difference in performance between a more referential condition versus a more straightforward condition like the visible-hiding conditions. Children, conversely, seem capable of appreciating the distinction between visible and invisible hiding for direct conditions. Children were not successful using video in this experiment, so no interpretation can be made about performance in visible versus invisible hiding for these conditions.

The next two experiments explore further what might be at the origin of the lesser performance by human children in the context of video conditions in the present task. This is accomplished by trying to boost this performance by either providing them with enhanced perceptual cues regarding the hiding objects (Experiment 2) or, in addition, providing them with these enhanced perceptual cues as well as enhanced verbal instructions (Experiment 3). In other words, we explored what might help 3-year-old children to equate and eventually bypass the chimpanzees' and at least one of the orangutans' performance in all of the hiding conditions tested here.

Experiment 2 was meant to test the possibility that in Experiment 1 children failed to discriminate the two hiding containers. To highlight perceptual contrast among the hiding containers, we replaced the original tubes with a green watering can, a red cup, and a blue rectangular box. The basic rationale for Experiment 2

was that if perceptual discrimination was a factor, these new objects should help improve children's search performance.

Experiment 2

Method

Participants. Twelve 36-month-old children ($M = 36.6$ months, $SD = 1.1$; 7 boys and 5 girls) participated. They were recruited from a participant pool consisting of children from the Atlanta area. Parents were contacted by phone and invited to participate in the study. Participants were mostly White and middle class. They received a small gift for their participation. Parents were allowed to remain present during all phases of testing. The children were tested at the Child Study Center at Emory University. Signed consent forms were obtained from parents of all children. We tested 3 additional children who were excluded because of fatigue and fussiness.

Apparatus and procedure. Hiding objects were three different and highly contrasted toy artifacts: a green plastic watering can, approximately 6 in. \times 6 in. (15.2 cm \times 15.2 cm); a red plastic swirly cup, approximately 4 in. \times 6 in. (10.2 cm \times 15.2 cm); and a blue plastic rectangular box, approximately 5 in. \times 6 in. (12.7 cm \times 15.2 cm). As in Experiment 1, the object hidden was a small stuffed bear, approximately 1 in. \times 2.5 in. (2.5 cm \times 6.4 cm). The procedure was the same as in Experiment 1.

Coding and dependent measures. Coding and dependent measures were the same as in Experiment 1. Percentage agreement between coders on correct and incorrect responses of all tested participants across all trials in each condition was above 99%.

Results

We analyzed the mean percentage of correct responses, using a two-way, repeated-measures ANOVA with hiding (visible vs. invisible) and video (video vs. direct) as within-subjects variables. The ANOVA revealed a main effect of hiding, $F(1, 11) = 11.80$, $p < .01$ (visible hiding: $M = .76$, $SD = .05$; invisible hiding: $M = .59$, $SD = .05$). This indicates that performance in the visible-hiding conditions was significantly different than performance in the invisible-hiding conditions. The analyses did not reveal a main effect of direct versus video conditions nor any significant interaction between this variable and the hiding variable.

As shown in Figure 4, sign tests revealed that children in Experiment 2 performed significantly above chance level in both the direct, visible-hiding and video, visible-hiding conditions ($p < 4.64 \times 10^{-10}$ and $p < .006$, respectively). However, the analysis revealed that children were not significantly above chance in either

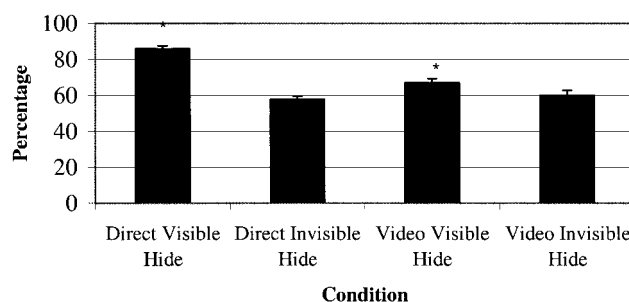


Figure 4. Mean percentage of correct responses in Experiment 2 as a function of condition for 3-year-old children using new containers. Error bars reflect standard error as a function of mean percent correct. * $p < .05$.

the direct, invisible-hiding condition or the video, invisible-hiding condition.

Discussion

The results of Experiment 2 support our prediction that increasing the perceptual distinctiveness between hiding locations would facilitate children's performance. In contrast to what we found in Experiment 1, the results illustrated that children performed significantly above chance in the direct, visible-hiding and video, visible-hiding conditions as a result of the use of the more distinctive containers. These more dissimilar hiding locations appeared to help children in locating the hidden object in the video, visible-hiding condition.

Performance for children was not, however, significantly above chance in either the direct, invisible-hiding or video, invisible-hiding conditions. This is not surprising given the results of Experiment 1 and our view that these invisible hiding conditions may be too vague for children in the absence of verbal instructions from the adult experimenter.

The third experiment provided additional verbal cues in these conditions. The rationale was that perhaps such additional cues would enhance 3-year-olds' performance to become more comparable to the performance of chimpanzees. In other words, we hypothesized that children's performance is highly dependent on verbal cues. We tested further this idea in the last experiment.

Experiment 3

Method

Participants and apparatus. Thirteen 36-month-old children ($M = 36.4$ months, $SD = 1.4$; 7 boys and 6 girls) participated. They were recruited from a participant pool consisting of children from the Atlanta area. Parents were contacted by phone and invited to participate in the study. Participants were mostly White and middle class. They received a small gift for their participation. Parents were allowed to remain present during all phases of testing. The children were tested at the Child Study Center at Emory University. Signed consent forms were obtained from parents of all children. An additional 6 children were tested but excluded—5 for fatigue and fussiness, 1 for not following instructions. The apparatus in this experiment was the same as in Experiment 2.

Procedure, coding, and dependent measures. The procedure was the same as in Experiment 2 except that in both the direct, invisible-hiding and video, invisible-hiding conditions, children received additional verbal instructions. In particular, when shaking the appropriate hiding location, the experimenter also whispered, "This is where Mr. Bear is going to hide." This additional cue was whispered to stand for an additional piece of verbal information. The coding and dependent measures were the same in this experiment as in Experiment 2.

Results

We analyzed the mean percentage of correct responses using a two-way, repeated-measures ANOVA with hiding (visible vs. invisible) and video (video vs. direct) as within-subjects variables. The ANOVA revealed a main effect of video on 3-year-olds' performance, $F(1, 12) = 7.58$, $p < .05$ (video: $M = .71$, $SD = .05$; direct: $M = .84$, $SD = .04$). Children's performance in the direct condition was significantly different compared with the video condition. In contrast to Experiment 2, the analyses did not reveal

a main effect of hiding. In addition, there was no significant interaction between hiding and video conditions.

As shown in Figure 5, sign tests revealed that children were significantly above chance in their choice in all four conditions (direct, visible hiding: $p < 4.31 \times 10^{-7}$; direct, invisible hiding: $p < 8.58 \times 10^{-9}$; video, visible hiding: $p < .0009$; video, invisible hiding: $p < .0001$).

Discussion

Results from Experiment 3 supported the original prediction that increased verbal instructions further enhance the performance of 3-year-old children in the direct, invisible and video, invisible-hiding conditions. In this last experiment, children performed significantly above chance in all four conditions. The results illustrate that increased verbal instruction during the invisible-hiding conditions was enough to boost 3-year-olds' performance, making it comparable to chimpanzees' performance.

General Discussion

In this research, we assessed apes and young children's ability to use video images as referential to real-world events. Reversing the typical comparative approach by which researchers adapt human tasks to test nonhuman subjects, we tested human children in a task that was first designed and adapted for the testing of apes. Within this context, we found that chimpanzees tended to fare better than children aged up to 3 years when using video cues. Overall, the 4 chimpanzees and at least 1 of the 2 orangutans that were tested demonstrated greater ability to use video cues to guide their search of a reward hidden in one of two locations. The chimpanzees' performance tended to be equally good whether the cues guiding their search were directly perceived or perceived on a TV screen. The chimpanzees clearly demonstrated an ability to (a) pick up relevant hiding information pertaining both to the hiding action of the experimenter and the particular physical characteristics of the hiding location (i.e., distinctive features of the tube in which the reward was concealed); (b) store this perceptual information for a short period (approximately 5 s); (c) retrieve it during the actual search task when both hiding tubes

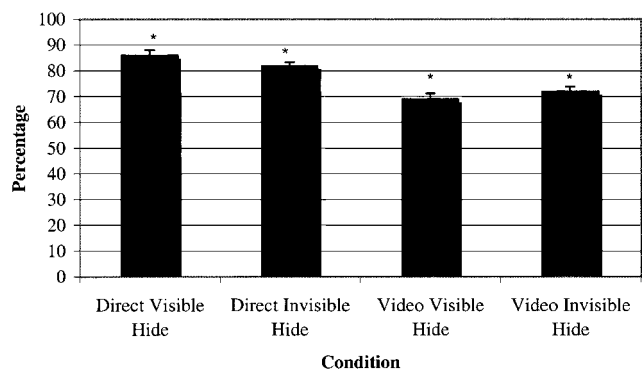


Figure 5. Mean percentage of correct responses in Experiment 3 as a function of condition for 3-year-old children with increased verbal cues. Error bars reflect standard error as a function of mean percent correct. * $p < .05$.

were presented within reach for choice; and (d) transpose and map video information onto the real space in which the search took place.

In sharp contrast, children up to 3 years of age demonstrated a marked deterioration of their performance when provided with video cues on the hiding location of the reward. In comparable procedural circumstances, children express a difficulty in transposing and mapping video information onto the real space of the search. Existing developmental research suggests that by 30 months (2.5 years) of age, children begin to show some ability to use video cues to guide behavior in a real-world setting (Troseth & DeLoache, 1998). This is the reason why our first hypothesis was that 30-month-old children should have been comparable to chimpanzees in their successful ability to use video cues to guide their search for a reward (e.g., Menzel et al., 1978). Our results from Experiment 1 do not confirm this hypothesis, showing the greater ability of chimpanzees compared with children up to 3 years of age. There are at least three possible accounts of such discrepancy in performance.

The first possibility is that it is only beyond 3 years of age that children can flexibly use video or real cues interchangeably in the absence of other verbal and perceptual cues to guide their action in a real setting, an ability expressed by the chimpanzees we tested here. A closer look at the developmental research on the use of video cues to search for objects indicates that children up to 3 years of age are not totally incapable of detecting and using video cues to guide their search. Instructing 2-year-olds that the TV is actually a window through which they perceive real events helps children in their subsequent search of a hidden object (Troseth & DeLoache, 1998). Schmitt and Anderson (2002) also report that in 2-year-old children, their first search trial for a toy in a real setting based on video cues is typically above chance, eventually deteriorating in subsequent trials because of perseveration errors with the child searching at the location where the toy was found last. Therefore, the ability to use video cues to guide behavior is not merely absent in toddlers but seems to become more reliable between 2 and 3 years of age.

A second possibility is that the lesser performance of children is perceptual and is thus due to children's difficulty in discriminating between the hiding objects (the tubes specifically designed for the testing of apes in the present experiment). Experiment 2, which tested for this possibility by introducing children with contrasted hiding containers, demonstrated that this assumption may have merit. With the use of more contrasted hiding containers, otherwise following the same procedure as in Experiment 1, 36-month-olds showed comparable above chance search performance in both the direct, visible-hiding and video, visible-hiding conditions. Therefore, children, in contrast to chimpanzees, appeared to need enhanced perceptual cues to be as successful in the direct and video conditions. This perceptual enhancement seems to help children in relating what is seen on the TV and what is subsequently presented in the real setting for the search test. We conclude that children up to 3 years of age are less flexible in using video cues to guide their search compared with chimpanzees. One could argue that this discrepancy could be motivational because for chimpanzees we used a food reward. In addition, children were rewarded in the form of praise regardless of a correct or incorrect choice. Although it is difficult to argue against this possibility, we found that children were in general enticed into searching for the

toy, enjoying the game as a novel play activity, in a new environment with new people. Providing children with food for a search reward—in particular, candy, as it is the universal kid's reward and equivalent to the treat bananas represent for apes—was not feasible because of ethical and health reasons. However, we are fairly confident that apes and children were both as enticed to search for the reward, whether it was a food or nonfood item.

A third possibility was that children might show an overall lesser performance when compared with chimpanzees in an experiment primarily designed for the latter because of the lack of verbal support as the trademark of a human environment. In other words, in these strict comparable circumstances, children would be at a disadvantage, tested in a nonverbal task that would be ecologically invalid for them. This possibility does not seem to be feasible in relation to the lesser performance in the video conditions because Experiment 2 showed that merely replacing the hiding tubes with more contrasted objects boosted 3-year-olds' performance using the same, ape-adapted, nonverbal procedure. However, Experiment 3 demonstrated that providing verbal instructions and support does indeed have an impact on 3-year-olds' search, boosting their performance in the invisible-hiding conditions, whether direct or video.

In all, the results show that the performance of 3-year-old children does improve significantly depending on both the perceptual (Experiment 2) and verbal cue enhancement (Experiment 3) provided by the experimenter during the manipulation. These supplemental cues facilitated increased performance in 3-year-old children and point to the difference in cognitive ecology between human and nonhuman primates. This result underlies the importance of factoring such differences in any comparative cognition research as well as the need to explore the nature of such differences.

Finally, we expected that apes, in general, would rely more on stimulus enhancement cues in their search (i.e., a preference for the target object that was previously manipulated more by the experimenter). This would mean that in comparison to children, apes would rely more on surface stimulation, their choice merely based on the relative amount of movement previously associated with either hiding object location. We reasoned that if this was correct, then apes compared with children should show no difference in performance between visible-hiding and invisible-hiding conditions, given that in both conditions, the baited object was manipulated more by the experimenter (see introduction). Because the apes tended to show increased performance in their search in the visible-hiding compared with the invisible-hiding condition, we did not confirm this hypothesis. In fact, apes seemed to have greater difficulty with invisible-hiding conditions. Therefore, what guided apes' search appears to have been more than stimulus enhancement. In other words, the fact that the invisible-hiding condition corresponded to lesser performance indicates that apes based their search on different cues regarding the hiding event, not the mere presence or absence of movements associated with a particular object during the cue phase of the experiment. More research is needed to capture the exact nature of the meaningful cues apes are shown to be capable of using to guide their behavior, whether these cues are presented in the real setting or by means of a TV.

In conclusion, the present experiments reveal remarkable abilities in apes in tasks that are designed to tap into their cognitive

ecology (i.e., simple nonverbal tasks involving immediate reward). When comparing human children within the same circumstances, we find surprising discrepancies that tend to be reversed to what is typically described in the comparative literature: Chimpanzees tend to fare better than children because they do not appear to rely on verbal cues and other symbolic instructions like children do. The phenomenon reported here points to the necessity for comparative ape research to consider the particular social and cognitive ecology of the compared species.

References

- Call, J., & Rochat, P. (1996). Liquid conservation in orangutans (*Pongo pygmaeus*) and humans (*Homo sapiens*): Individual differences and cognitive strategies. *Journal of Comparative Psychology*, 110, 219–232.
- Call, J., & Rochat, P. (1997). Perceptual strategies in the elimination of physical quantities by orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology*, 111, 315–329.
- Callaghan, T. C. (2000). Factors affecting children's graphic symbol use in the third year: Language, similarity, and iconicity. *Cognitive Development*, 15, 185–214.
- DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. *Child Development*, 62, 736–752.
- DeLoache, J. S. (2000). Dual representation and young children's use of scale models. *Child Development*, 71, 329–338.
- Fouts, R. (1998). *Next of kin: What chimpanzees have taught me about who we are*. New York: Morrow & William.
- Gardner, R. A., Gardner, B. T., & Van Cantfort, T. E. (Eds.). (1989). *Teaching sign language to chimpanzees*. Albany: State University of New York Press.
- Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language & thought: Interrelations in development* (pp. 225–277). Cambridge, England: Cambridge University Press.
- Hallock, M. B., & Worobey, J. (1984). Cognitive development in chimpanzee infants (*Pan troglodytes*). *Journal of Human Evolution*, 13, 441–447.
- Hopkins, W. D., Fernandez-Carriba, S., Wesley, M. J., Hostetter, A., Pilcher, D., & Poss, S. R. (2001). The use of bouts and frequencies in the evaluation of hand preferences for a coordinated bimanual task in chimpanzees (*Pan troglodytes*): An empirical study comparing two different indices of laterality. *Journal of Comparative Psychology*, 115, 294–299.
- Kuhlmeier, V. A., & Boysen, S. T. (2001). The effect of response contingencies on scale model task performance by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 115, 300–306.
- Kuhlmeier, V. A., & Boysen, S. T. (2002). Chimpanzees (*Pan troglodytes*) recognize spatial and object correspondences between a scale model and its referent. *Psychological Science*, 13, 60–63.
- Kuhlmeier, V. A., Boysen, S. T., & Mukobi, K. L. (1999). Scale-model comprehension by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 113, 396–402.
- Mathieu, M., & Bergeron, G. (1981). Piagetian assessment on cognitive development in chimpanzees (*Pan troglodytes*). In A. B. Chiarelli & R. S. Corruccini (Eds.), *Primate behavior and sociobiology* (pp. 142–147). Berlin, Germany: Springer-Verlag.
- Menzel, E. W., Jr. (1971). Communicating about the environment in a group of young chimpanzees. *Folia Primatologica*, 15, 220–232.
- Menzel, E. W., Jr., Premack, D., & Woodruff, G. (1978). Map reading by chimpanzees. *Folia Primatologica*, 29, 241–249.
- Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development*, 69, 295–308.
- Natale, F. (1989). Stage 5 object-concept. In F. Antinucci (Ed.), *Cognitive structure and development in nonhuman primates* (pp. 97–112). Hillsdale, NJ: Erlbaum.
- Premack, D. (1986). *Gavagai! The future history of the animal language controversy*. Cambridge, MA: MIT Press.
- Redshaw, M. (1978). Cognitive development in human and gorilla infants. *Journal of Human Evolution*, 7, 133–141.
- Rochat, P. (2001). Origins of self-concept. In G. Bremner & A. Fogel (Eds.), *Blackwell handbook of infant development. Handbooks of developmental psychology* (pp. 191–212). Malden, MA: Blackwell.
- Savage-Rumbaugh, E. S., McDonald, K., Sevcik, R. A., Hopkins, W. D., & Rupert, E. (1986). Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (*Pan paniscus*). *Journal of Experimental Psychology: General*, 115, 211–235.
- Schmitt, K. L., & Anderson, D. R. (2002). Television and reality: Toddlers' use of visual information from video to guide behavior. *Media Psychology*, 4, 51–76.
- Seidenberg, M. S., & Petitto, L. A. (1987). Communication, symbolic communication, and language: Comment on Savage-Rumbaugh, McDonald, Sevcik, Hopkins, and Rupert (1986). *Journal of Experimental Psychology: General*, 116, 279–287.
- Striano, T., Tomasello, M., & Rochat, P. (2001). Social and object support for early symbolic play. *Developmental Science*, 4, 442–455.
- Thorpe, W. H. (1956). *Learning and instinct in animals*. London: Methuen.
- Tomasello, M., & Call, J. (1997). *Primate cognition*. New York: Oxford University Press.
- Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development*, 68, 1067–1080.
- Tomasello, M., Striano, T., & Rochat, P. (1999). Do young children use objects as symbols? *British Journal of Developmental Psychology*, 17, 563–584.
- Troseth, G. L., & DeLoache, J. S. (1998). The medium can obscure the message: Young children's understanding of video. *Child Development*, 69, 950–965.
- Vauclair, J. (1990). Primate cognition: From representation to language. In S. T. Parker & K. R. Gibson (Eds.), *Language and intelligence in monkeys and apes* (pp. 312–329). Cambridge, England: Cambridge University Press.

Received April 8, 2002

Revision received April 14, 2003

Accepted April 19, 2003 ■