Developmental Science

Developmental Science 11:5 (2008), pp 778–792 DOI: 10.1111/j.1467-7687.2008.00729.x

PAPER

How young children and chimpanzees (*Pan troglodytes*) perceive objects in a 2D display: putting an assumption to the test

Katherine A. Leighty,[1] Charles R. Menzel[2] and Dorothy M. Fragaszy[3]

1. *Education and Science, Animal Programs Administration, Disney's Animal Kingdom, USA*
2. *Language Research Center, Georgia State University, USA*
3. *Department of Psychology, University of Georgia, USA*

Abstract

*Object recognition research is typically conducted using 2D stimuli in lieu of 3D objects. This study investigated the amount and complexity of knowledge gained from 2D stimuli in adult chimpanzees (*Pan troglodytes*) and young children (aged 3 and 4 years) using a titrated series of cross-dimensional search tasks. Results indicate that 3-year-old children utilize a response rule guided by local features to solve cross-dimensional tasks. Four-year-old toddlers and adult chimpanzees use information about object form and compositional structure from a 2D image to guide their search in three dimensions. Findings have specific implications to research conducted in object recognition/perception and broad relevance to all areas of research and daily living that incorporate 2D displays.*

Introduction

Researchers in the field of visual perception are interested in how observers recognize objects, and are attempting to determine the shared mechanisms underlying this process across species (Rilling, LaClaire & Warner, 1993). To assess components of the recognition process, researchers typically present participants with two-dimensional (2D) depictions of three-dimensional (3D) objects (Bovet & Vauclair, 2000). Experiments are conducted in this fashion due to the ease of manipulating stimulus form, rotation, randomization, viewpoint, and ease of data collection. Additionally, tests of the visualization of objects in space from a 2D depiction (i.e. paper folding, mental rotation, object reconstruction from parts or views) are thought to require the participant to form a 3D mental representation of the 2D depiction to solve the task. These forms of research are based on the assumption that cognitive processing of a stimulus presented in two and three dimensions results in perceptual experiences that can be equated (Reid & Spetch, 1998; Shyi & Huang, 1995).

Fagot, Martin-Malivel and Dépy (2000) have proposed that perceivers may use any or all of three modes of processing the relationship between an object and its 2D depiction. The first mode is termed the independence mode. Here the processing of the 2D depiction is unrelated to the representational content, thus the processing of the object and its 2D depiction are independent. The second mode of processing is the confusion mode. In this mode the depiction and object are processed in the same way because the subject does not differentiate between them. That is, the image and the object are perceived to be the same. The third and final mode is the equivalence mode. Here, the image and object are seen to be symbolically related but are physically differentiated. It is in this mode that the subject comprehends that the 2D image functions to depict a 3D object (Fagot *et al.*, 2000).

We propose two submodes of the equivalence mode; featural and complex. In *featural equivalence processing*, the observer uses local features to guide responding. That is, the observer detects a feature in one dimension and matches it to that feature in the other dimension (e.g. a facial feature, the round one, the blue item, the one with the red stripe). Thus, the subject may only need to represent a particular local feature across dimensions and not the object's form. In *complex equivalence processing*, knowledge of the 3D object's global form as well as its compositional structure is gained from the 2D image. Subjects operating in this submode do not simply match for features across dimensions. Instead, they can recognize relational elements within the object across dimensions (e.g. lower right corner, bottom middle drawer). The degree of object complexity that can be used by this mental process must be determined and is assumed to vary developmentally and across species.

Address for correspondence: Katherine A. Leighty, Animal Programs Administration, Disney's Animal Kingdom, P.O. Box 10000, Lake Buena Vista, FL 32830, USA; e-mail: Katherine.Leighty@Disney.com

Most adult humans can recognize objects from their 2D depictions and differentiate between them as in the equivalence mode (Davenport & Rogers, 1971; Fagot *et al.*, 2000; Watanabe, 1997). Yet, there has been debate in the literature as to the necessity of experience to recognize 3D objects from 2D depictions. Cross-cultural comparisons have demonstrated that participants with little to no experience viewing 2D images were often unable to recognize the objects depicted. It was only with repeated exposure and demonstration of pertinent features that participants came to perceive the images using the equivalence mode (Miller, 1973; Deregowski, Muldrow & Muldrow, 1972). Because experience is clearly a mediating factor in the perception of 2D depictions, investigating these abilities in young children and nonhuman subjects is a scientific necessity before interpreting results of the numerous studies conducted with these subjects utilizing 2D stimuli and extrapolating the findings to 3D perception.

Using visual preference and habituation procedures, Slater, Rose and Morison (1984) examined the perception of 2D depictions of objects in human infants. Their findings indicated that newborns have the ability to take the first step in processing the relationship between the 2D depictions and 3D objects in that they can discriminate between the two forms. To address the issue of cross-dimensional equivalence in 5- and 6-month-old infants, Rose (1977) used a visual fixation paradigm with two objects, a sunburst and a diamond constructed of black wooden strips .6 cm thick, along with black and white photographs of each. Findings indicated that infants could detect the difference in dimensionality and were able to transfer information across dimensions. Yet, because only two stimuli were used, and because they were quite different in structure, infants may have been matching local features (i.e. featural equivalence submode) and not representing the objects' structure across dimensions (i.e. complex equivalence submode).

Further, an interesting phenomenon has been documented in 9-month-old infants (DeLoache, Pierroutsakos, Uttal, Rosengren & Gottlieb, 1998; Pierroutsakos & DeLoache, 2003). Infants at this age will often manually investigate 2D depictions of 3D objects. This behavior seems to suggest some degree of processing of 2D images using the confusion mode. The investigators suggest infants' knowledge of the symbolic nature of 2D images lags behind the ability to perceive the pictured information. That is, they are uncertain as to the nature of pictures. While the behavior directed to the 2D images is more tentative than that directed to the 3D object, it does suggest at least a mild degree of confusion.

Perhaps most important to the question of cross-dimensional perception is how and when children learn to utilize the information in a 2D display to direct actions in the 3D world. Brown (1969) sought to begin to answer this question by addressing whether children can align 3D objects to match a depiction of their layout presented in two-dimensions. He classified the age-related differences and constructed five stages of comprehension of the 2D depiction. He found that it was not until the age of 6 that children demonstrated complete understanding of the symbolic nature of photographs by correctly arranging items to match the photograph, including matching item placement in depth.

DeLoache and colleagues have conducted many studies using a cross-dimensional search paradigm to address scene perception in young children. By presenting young children with pictures of an object hidden in a room and then introducing the child into the depicted environment, they were able to document the participants' search patterns. Children 24 months of age were unable to find the hidden toy, while 30-month-old children were successful. They believe that these findings indicate that it is not until 30 months of age that children can comprehend the function of 2D images of representing a 'current reality' (DeLoache & Burns, 1994). These findings are comparable to those of Troseth and DeLoache (1998) using video instead of photographic 2D images. In this study, children again searched for a hidden object, but this time they were shown a video of the object being hidden in the room. It was found that 24-month-olds could not complete the task, while 30-month-olds were successful. The findings of these studies of DeLoache and colleagues do provide evidence of comprehension of the representational function of 2D images. Yet, one must consider the possibility that children solve these tasks by using featural equivalence. That is, 30-month-old children may be guided in their search by a relevant feature of the scene present in both the 2D image and 3D environment (i.e. search under the chair). It remains unclear as to whether children at this age have developed the ability to process 2D images in the proposed complex equivalence submode.

These findings in young children beg the question of what mode of processing of 2D stimuli is utilized in nonhuman subjects. Nonhuman primates are widely studied in most areas of visual perception research, especially neural imaging studies that are meant to allow us to comprehend human functioning by studying brain–behavior relations in nonhuman primates. The vast majority of these testing paradigms present subjects with 2D depictions of 3D objects and many extrapolate the results to humans. Therefore, it seems that research to understand if and how primates perceive 2D depictions of 3D objects is a necessity that has often been neglected (Fagot *et al.*, 2000). Further, the handful of studies that have attempted to parse out the perceptual processing of 2D stimuli in nonhuman primates have produced contradictory findings.

Cross-modal matching-to-sample (MTS) is one technique that was used in early studies of this cross-dimensional perception in nonhumans. Davenport and Rogers (1971) presented two chimpanzees and one orangutan with a photograph and had subjects select the appropriate match to the item in the photograph by haptically (in the absence of vision) investigating a pair of objects. The 3D objects were described as 'highly distinguishable' such as a tap handle, padlock, fishing lure, and bobbin. It was found that the apes correctly matched the 3D object to the 2D depiction at levels significantly above chance. Malone, Tolan and Rogers (1980) conducted a similar study with two male rhesus macaques and found that both monkeys learned to complete the task with extensive practice. In contrast, Winner and Ettlinger (1979) failed to replicate the findings of Davenport and Rogers (1971) with two juvenile chimpanzees. Thus, previous results using cross-modal techniques indicate that some, but not all, apes and macaques learn to use photographs as discriminative stimuli to select 3D objects haptically, but the findings do not specify which aspects of the photographic stimuli were crucial to correct performance.

Other techniques that are not cross-modal in nature have also been utilized to examine abilities for cross-dimensional transfer in nonhumans. Zimmerman and Hochberg (1970, 1971) trained infant rhesus macaques to discriminate between flat and solid objects, such as wooden squares and cubes. After learning this discrimination, the infants transferred the discrimination to photographs of these objects (Zimmerman & Hochberg, 1970, 1971). These findings are difficult to interpret due to the extreme physical contrast that existed between the 3D objects. Subjects may have used the featural equivalence submode to guide their actions rather than the complex equivalence submode.

Utilizing a similar methodology, Martin-Malivel (1998) examined the abilities of baboons to distinguish between objects and their 2D depictions using a go/no-go discrimination task. Subjects had low rates of success in completing the cross-dimensional transfer phase of this task and in many cases positive transfer was not achieved. Therefore, the equivalence mode of processing was clearly not available to all subjects in all conditions. An interesting finding of a facilitating effect of object familiarity in subjects able to complete this task supports the notion that familiarity affords processing of 2D images in the equivalence mode.

Perhaps the most robust results have been obtained using a categorization paradigm in which animals are not matching pictures of objects to the objects themselves but rather are classifying them as a specific type of object. Several studies have found that chimpanzees have the capacity to match pictures of objects to their assigned categories (e.g. 'food/non-food', 'tool/food'; Premack & Woodruff, 1978; Savage-Rumbaugh, Rumbaugh, Smith & Lawson, 1980; Tanaka, 1996). Similar abilities of transferring a learned categorical discrimination across dimensions have been observed in olive baboons (Bovet & Vauclair, 1998). Results from cross-dimensional categorization tasks such as these lend stronger support to the use of the equivalence mode of processing 2D stimuli, yet they do not rule out the use of the featural equivalence submode since features can be used to define membership in a category. Martin-Malivel and Fagot (2001) presented an interesting variant of this methodology using a cross-modal go/no-go categorization task in which baboons were presented with images of humans or baboons prior to discriminating between human and baboon vocalizations. One baboon demonstrated a cross-modal priming effect of reduced reaction times when the picture and vocalization were conceptually related. This effect was replicated when the images were degraded, thus reducing the likelihood that object features were the sole basis of the priming effect.

A number of studies have demonstrated species-appropriate responses to 2D depictions of biologically relevant stimuli. For example, von Heusser (as cited in Bovet & Vauclair, 2000) presented a marmoset with photographs of butterflies and snakes. He noted the production of a grasping response to butterfly images and fear reactions to the snakes. Bovet and Vauclair (1998) observed baboons reaching to 2D cutouts of food images but not to non-food images. Behavioral responses such as these to 2D stimuli suggest that processing may be occurring in the confusion mode. Using comparable methods, Sackett (1965, 1966) and Rosenfeld and van Hoesen (1979) presented slides of different social situations and partners to rhesus monkeys reared in isolation and observed appropriate fear responses to threats. Similar results were obtained when cynomolgus monkeys were shown slides depicting threatening individuals (Kyes, Mayer & Bunnell, 1992). Analogous results have been obtained when animals are presented with dynamic video images. Juvenile bonnet macaques demonstrate species-appropriate responses when they observe video of individual social displays (Plimpton, Swartz & Rosenblum, 1981). Additionally, squirrel monkeys produce species-appropriate alarm calls and threat reactions when shown video of different predators (Herzog & Hopf, 1986). Finally, Boysen and Berntson (1986) documented elevated heart rates in chimpanzees after viewing photographs of familiar keepers and agonistic conspecifics.

Evidence from nonhuman primates suggests that both apes and monkeys respond to 2D depictions much as they respond to 3D objects to some degree. Unfortunately, it is not clear on what level they perceive the relationship between objects and 2D images. Do nonhuman primates always distinguish between objects and images (i.e. they do not operate in the confusion mode), and if so, do they utilize a response rule based on local features (i.e. featural equivalence submode), or do they gain knowledge of the structure of the object (i.e. complex equivalence submode) from the 2D depiction?

The methodology used in this study was derived from that which has been used to assess the abilities of young children and chimpanzees to use scale-models, pictures, and video presentations of real-world situations. In these studies the subjects are presented with information regarding the location of a hidden item in a miniature 3D (scale-model) or 2D (picture, video) format. Subjects are then allowed to search the 3D environment where the object is hidden and their search behavior is documented. It is not until children reach 30–36 months of age that they can successfully use a model or 2D image

Table 1 *Chimpanzee subjects' background information*

Subject	Dob	Lexigram training	2D experience
Lana	10/7/1970	Yes	Exposure to television, videos and photographs Computerized tasks (Hopkins, Washburn & Hyatt, 1996; Rumbaugh, Washburn, Savage-Rumbaugh & Hopkins, 1991; Rumbaugh, 1977)
Sherman	5/10/1973	Yes	Exposure to television, videos and photographs Live video interactions Computerized tasks (Hopkins *et al.*, 1996; Rumbaugh *et al.*, 1991; Savage-Rumbaugh, 1986; Menzel, 1985)
Panzee	12/31/1985	Yes	Exposure to television, videos and photographs Computerized tasks 2D to 3D transfer tasks (Hopkins *et al.*, 1996; Rumbaugh *et al.*, 1991; Brakke & Savage-Rumbaugh, 1995, 1996; Menzel, 2005)
Mercury	11/15/1986	No	Exposure to television and videos Computerized tasks (Hopkins *et al.*, 1996; Rumbaugh *et al.*, 1991)

to guide their search (DeLoache, 2000; DeLoache & Burns, 1994; Troseth & DeLoache, 1998). When presented with a 3D scale-model or its 2D video image, chimpanzees demonstrated the ability to use the information provided to guide their search when other strategies were eliminated or made ineffective (Kuhlmeier & Boysen, 2001; Kuhlmeier, Boysen & Mukobi, 1999). While compelling, these scale-model studies do not eliminate the potential use of featural matching to solve the task. That is, both young children and chimpanzees may search for the hidden item using distinctive local features (e.g. look under the round item, in the black box).

Chimpanzees show excellent spatial memory in large-scale outdoor environments. In studies of map reading and memory, four juvenile chimpanzees (Menzel, Premack & Woodruff, 1978) and an adult chimpanzee (Menzel, 2001, 2005) used a video representation of a field or a forest as a guide to locating objects in that area. The adult chimpanzee performed accurately with overnight delays exceeding 15 hours. The visual features that chimpanzees use to encode locations in complex naturalistic environments are not completely understood at this time, but the accuracy of the chimpanzees in video tasks encouraged the use of video presentation in the current study (see also Poss & Rochart, 2003).

In order to determine the level of processing of 2D images, we utilized a bottom-up approach and began by addressing the perception of individual objects across dimensions. In the pair of experiments presented here, the perceptual processing of 2D stimuli and comprehension of their relationship to 3D objects by chimpanzees and young children was examined using a titrated series of cross-dimensional search tasks. Systematic removal of discriminative local features from the test objects followed by ordered increases in their structural complexity across testing phases allowed for clear conclusions as to the mode of perceptual processing and the level of object complexity that can be represented across dimensions by these subjects. If 2D depictions are associated to 3D objects using the featural equivalence submode, performance on the cross-dimensional search task would decline with the removal of distinctive local features. If processing occurred using the complex equivalence submode, performance would not be affected by the removal of local features and increases in structural complexity in the absence of distinctive local features could be assessed.

Experiment 1

Method

Subjects

Four adult chimpanzees (aged 16 to 33) served as the subjects for this study (two male, two female). Subjects were housed and testing occurred at the Language Research Center of Georgia State University. All subjects have experience with 2D stimulus presentations of pictures and video, have participated in a variety of cognitive tests, and three of the four subjects have extensive training in the use of a lexigram system (see Table 1 for detailed training history). Normal diet and husbandry conditions were not altered during the course of this study.

Materials

The study was conducted in an indoor enclosure measuring 4.67 m × 4.01 m × 2.59 m. The front panel of the enclosure contained two inset testing stations, each measuring 48.26 cm × 91.44 cm × 58.42 cm and located 93.98 cm from the ground and 1.02 m apart (see Figure 1). An opaque panel measuring 1.50 m × 1.19 m was used to obstruct the chimpanzee's view from one testing station to another. A ledge was positioned inside the entire cage front such that the subjects could be seated during the testing session. Live images used during testing were captured using a Panasonic Hi-8 Camera and relayed to a 26″ Sony WEGA Trinitron Color Television.

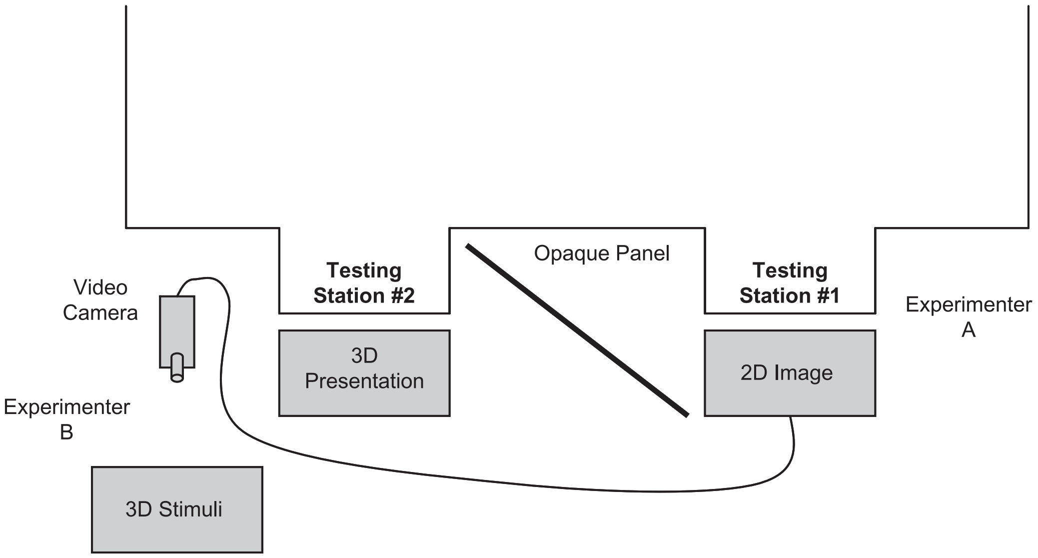

Figure 1 *Chimpanzee testing setup.*

Procedures

Two experimenters administered a titrated series of cross-dimensional search tasks to the subjects (see Figure 1). At the beginning of each trial the subject sat in front of a television screen positioned outside testing station 1 by experimenter 'A'. An opaque panel was positioned to block any viewing of testing station 2 and experimenter 'B' from testing station 1. Instead, the actions of experimenter 'B' were captured by a video camera and presented on the television screen. On each trial, experimenter 'B' would hide a marshmallow within an object on a table adjacent to testing station 2. This event was simultaneously viewed by the subject on the television screen. The screen was then turned off and experimenter 'B' would place the test stimuli at testing station 2. The subject was then led by experimenter 'A' past the opaque panel and seated at testing station 2. The subject was then permitted to touch one location and received the marshmallow if the correct location was selected.

To reduce experimenter bias, experimenter 'A' was not informed as to the location of the marshmallow across trials. Additionally, after placing the test items at station 2, experimenter 'B' would turn his/her back to the experimental setup to avoid providing any cues to the subject as to the marshmallow's location. Experimenter 'A' would then observe the subject's choice and call out the selected location. Experimenter 'B' would then announce whether the choice was of the 'correct' or 'incorrect' location. If the 'correct' location was selected on the first choice, experimenter 'A' would retrieve the marshmallow and hand it to the subject to consume. If the 'incorrect' location was selected, experimenter 'A' would show the subject the empty location followed by the 'correct' location. On each trial, experimenter 'B' recorded the first choice location and whether this choice was 'correct' or 'incorrect'.

Subjects were tested between two and four days per week for approximately 15 minutes per session (approximately 5–10 trials). A testing session was terminated when the subject no longer attended to the task and/or left the testing area. The criterion for completion of a testing phase was selecting the correct location of the hidden marshmallow on 17 of 20 consecutive trials (85% correct). Attainment of criterion performance on a testing phase advanced the subject to the next phase of testing. Testing was terminated when the subject achieved criterion on all phases or no longer attended and/or participated in a testing phase after completing a minimum 50 trials on that phase.

Eight titrated phases of testing were presented:

Phase 1: Color and Form MTS

In this first phase of testing, subjects learned the cross-dimensional search procedure by locating the hidden marshmallow beneath the correct sand mold (averaging 15 cm^3) in a two-choice match-to-sample (MTS) task (see Figure 2a). An array of six sand molds were used in this phase and the selection and location of all stimuli, in this phase as well as subsequent testing phases, were randomized across trials. During the MTS procedure, experimenter 'B' hid the marshmallow beneath a sand mold positioned on a tray while the subject viewed this event on the television screen. The screen would then be turned off and experimenter 'B' would add a second sand mold of a different shape and color as the distracter stimulus. Experimenter 'B' would then place the tray containing these two items (in the same spatial orientation as during baiting) at testing station 2 and would then turn away from the testing stimuli. Experimenter 'A' would then lead the subject from behind the opaque panel to testing station 2 and the subject would make a selection between the correct and distracter items. This cross-dimensional procedure required the subject to identify the correct item using properties of color and form.

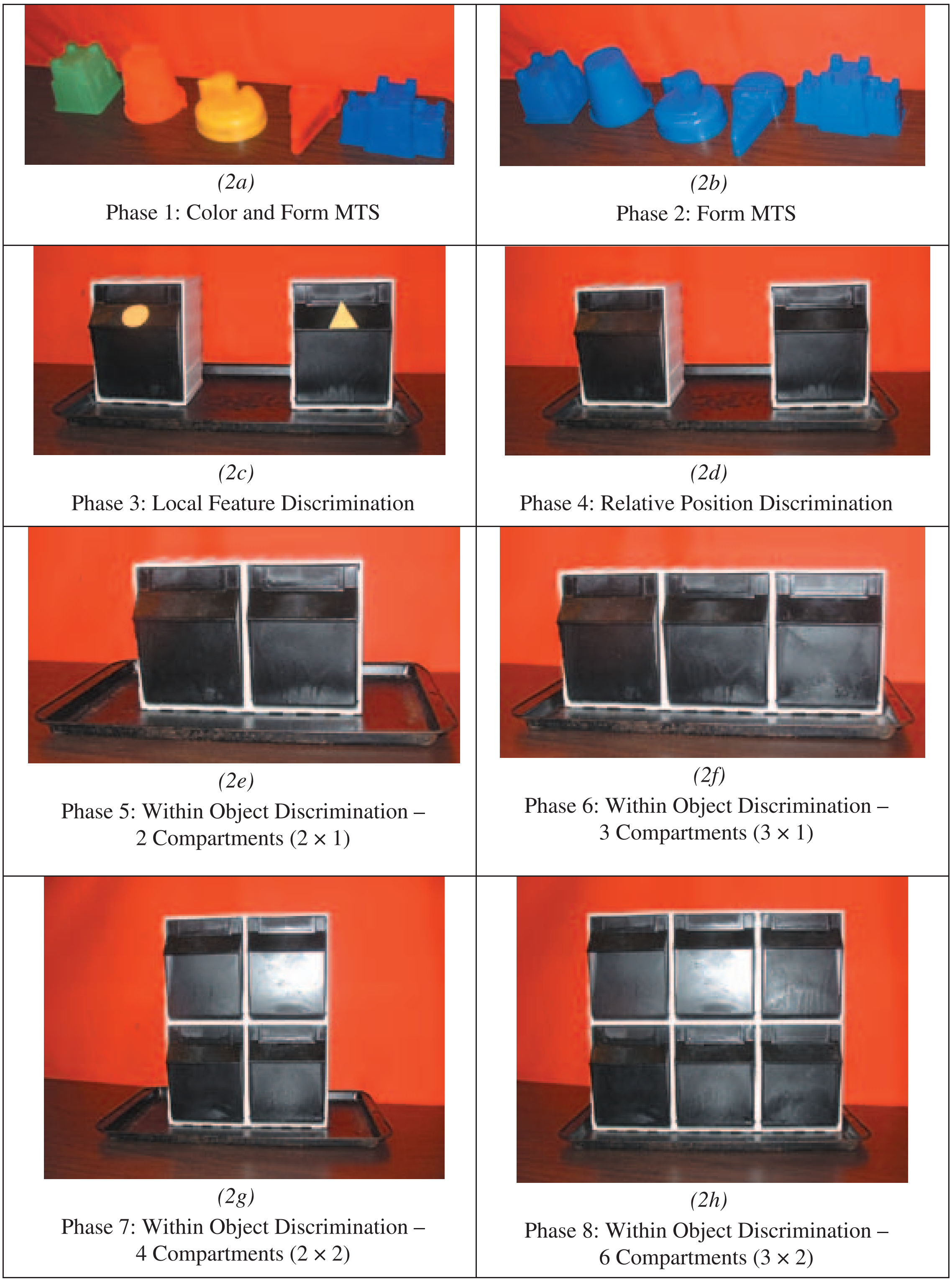

Figure 2 *Stimuli used in eight titrated phases of testing.*

Phase 2: Form MTS

The second phase of testing followed the same MTS procedure as the Color and Form MTS procedure (Phase 1) except that in this phase, the sand molds of varied shape were all painted blue (see Figure 2b). Controlling for color differences in this way required subjects to identify the correct item cross-dimensionally using properties of form alone.

Phase 3: Local Feature discrimination

In phases 3–8 the stimuli used to conceal the marshmallow were structurally identical tip-out bins (12 cm × 16 cm ×

11 cm). In this phase, two bins were used in a cross-dimensional discrimination procedure, each marked with a distinctive local feature (circle and triangle) (see Figure 2c). This testing phase required the subject to identify the location of the hidden marshmallow by recognizing the distinctive local feature or the container's relative position (left/right) across dimensions.

At the start of each trial, the camera was positioned on the two containers. Experimenter 'B' hid the marshmallow in one of the two containers and the subject viewed this event on the television screen. The screen was turned off and the procedure was carried out as in previous testing phases.

Phase 4: Relative position discrimination

In phase 4, two structurally identical containers were again used in a cross-dimensional discrimination procedure as in the previous testing phase. Here, the local features were removed and correct responses were mediated by the ability to represent the relative position (left/right) of two globally identical stimuli across dimensions (see Figure 2d).

Phase 5: Within Object discrimination – 2 compartments (2 × 1)

In this fifth phase of testing, the two identical containers were combined along their vertical axis to form a single object (see Figure 2e). Thus, the location of the hidden item in this phase had to be recognized across dimensions by its relative position (left/right) within the global form of a single object with two featurally identical compartments. All methods of presentation in this phase and in all subsequent phases were carried out in the same manner as in the previous testing phase.

Phase 6: Within Object discrimination – 3 compartments (3 × 1)

To increase the complexity of the cross-dimensional discrimination, three identical containers were combined to form a single object in this phase (see Figure 2f). Thus, to make a correct response, the subject had to discriminate between three featurally identical locations (left, middle, right) within the global form of a single object across dimensions.

Phase 7: Within Object discrimination – 4 compartments (2 × 2)

In this seventh testing phase, four containers were combined to form a single object with four compartments in a 2 × 2 layout (see Figure 2g). The subject therefore had to represent the correct location of the hidden marshmallow across dimensions by discriminating between four featurally identical compartments (top left/top right/bottom left/ bottom right) within the object's global form.

Phase 8: Within Object discrimination – 6 compartments (3 × 2)

In the final phase of testing, the complexity of the stimulus was again increased by combining six identical containers to form a single object with six compartments in a 3 × 2 layout (see Figure 2h). In order to reliably produce the correct response, the subject had to represent the location of the hidden item across dimensions by discriminating between six featurally identical compartments (top left/top middle/top right/bottom left/bottom middle/bottom right) within the object's global form.

3D control trials

If a subject did not attain criterion performance on a testing phase (1–8), 3D control trials were conducted to ensure that performance deficits were not due to memory or attentional failures. During these trials, the subject was presented with the phase of the task it could not complete in the cross-dimensional form. The subject directly viewed the experimenter 'B' hide the marshmallow in the test stimuli at testing station 1 (3D). The experimenter would then move the items to testing station 2 and the subject would follow and make its choice of the location of the hidden item (3D). All subjects attained criterion level performance (85%) in the 3D control tests in the first block of 20 trials presented.

Analysis

The total number of trials presented prior to attaining criterion performance in each cross-dimensional phase was determined for each subject. To examine the distribution of errors in the most complex form of the task, errors in Phase 8 (Within Object discrimination – 3 × 2) were classified according to the categories 'Previous Location', 'Horizontal' or 'Vertical'. An error was defined to be a 'Previous Location' error if the compartment selected was the location of the hidden marshmallow on the previous trial. A 'Horizontal' error occurred when the subject chose a compartment on the same horizontal row as the correct location and it was not a 'Previous Location' error. A 'Vertical' error occurred when the subject chose the compartment within the same vertical column as the correct location and it was not a 'Previous Location' error. The distribution of these errors was examined using a Chi-Square test.

Results

Phase 1 (Color and Form MTS)–Phase 5 (Within Object discrimination – 2 × 1)

All four subjects attained criterion performance on these phases of the cross-dimensional search task with the number of trials to attain criterion ranging across phases and subjects from 20 to 165 trials. Number of trials to attain criterion performance in each phase for each

Table 2 *Number of trials presented to attain criterion performance* by phase in chimpanzee subjects*

Phase	Subject				Mean (*SD*)
	Lana	Mercury	Panzee	Sherman	
1 Color and Form MTS	20	107	40	83	62.50 (39.64)
2 Form MTS	20	29	126	59	58.50 (47.99)
3 Local Feature discrimination	84	25	20	39	42.00 (29.13)
4 Relative Position discrimination	35	165	20	20	60.00 (70.36)
5 Within Object discrimination (2 × 1)	20	55	20	85	45.00 (31.36)
6 Within Object discrimination (3 × 1)	52	CNA	69	234	118.33 (100.53)
7 Within Object discrimination (2 × 2)	101	X	119	59	93.00 (30.79)
8 Within Object discrimination (3 × 2)	37	X	CNA	CNA	37 (N/A)

CNA = Criterion Not Achieved.
X = Phase Not Attempted.
* Criterion Performance = Correct on 17 of 20 consecutive trials.

subject along with a mean and standard deviation value across subjects by phase is presented in Table 2.

Phase 6: Within Object discrimination – 3 compartments (3 × 1)

Three of four subjects attained criterion performance on this cross-dimensional discrimination procedure in which three containers were combined to form a single object with three identical compartments. Mercury was dropped from testing due to lack of participation after the minimum 50 attempts. The mean number of trials to criterion across the three successful subjects was 118.33 (*SD* = 100.53) (see Table 2).

Phase 7: Within Object discrimination – 4 compartments (2 × 2)

In this phase, four identical containers were combined to form a single object arranged in a 2 × 2 configuration. All three subjects that attempted this phase attained criterion performance with a mean number of trials to criterion of 93.00 (*SD* = 30.79) (see Table 2).

Phase 8: Within Object discrimination – 6 compartments (3 × 2)

In this final testing phase, six identical containers were combined to form a single object arranged in a 3 × 2 configuration. One of the three subjects that attempted this phase attained criterion performance; Lana in 37 trials (see Table 2). Panzee and Sherman were dropped from testing after the minimum 50 attempts due to lack of participation/attention.

Error analysis of Phase 8

Lana committed eight total errors in the 37 trials needed to attain criterion performance. Of these eight errors, two were 'Previous Location' errors. Thus, 25% of Lana's errors may have been due to proactive interference effects. Of the remaining errors, Lana made four 'Horizontal' errors and one 'Vertical' error (see Table 3). Due to the 3 × 2 configuration, horizontal errors were twice as likely as vertical errors in this phase. Therefore, Lana's distribution of 'Horizontal' and 'Vertical' errors did not differ from that expected by chance (binomial proportions, p = .637). Sherman and Panzee made a total of 64 errors in the 100 trials they attempted (50 each) before testing was halted due to lack of participation/attention. Of these 64 errors, 11 were 'Previous Location' errors (17%). These subjects made 35 'Horizontal' errors' and 3 'Vertical' errors (see Table 3). Thus, the distribution of errors for these two subjects was significantly different than that expected by chance (χ^2 = 11.066, p = .0009).

Table 3 *Distribution of errors in Phase 8 – Within Object discrimination (3 × 2) for chimpanzee subjects*

Error type	Subject	
	Lana	Sherman & Panzee
Previous location	2	11
Horizontal	4	35
Vertical	1	3
Total errors	8	64
Total trials	37 (Criterion achieved)	100 (Criterion not achieved)

Discussion

Results supported the assumption of previous research utilizing 2D stimuli that chimpanzees do have the capacity to learn to equate their perception of objects presented in a 2D depiction with their 3D form. Previous research on cross-dimensional perception of individual objects demonstrated the capacity to match 3D items and their 2D depictions in nonhuman primates but did not eliminate the possibility that subjects were following a response rule of matching for local features across dimensions to solve the tasks (proposed featural equivalence submode). Additionally, prior work did not address the specific structural knowledge of the object that is gained from the 2D image. Here, by invoking a titrated testing series of cross-dimensional search tasks it was possible to illuminate the level of object complexity that chimpanzees can extract from a 2D image with and without the use of local features.

In the early phases of the experiment, all chimpanzees demonstrated the ability to complete matching-to-sample (MTS) tasks using information gained from a 2D television monitor to act upon 3D objects. The two phases of MTS testing (Color and Form MTS, Form MTS) were presented to address the ability of chimpanzees to recognize objects across dimensions. By removing color as a featural cue in Phase 2 (Form MTS), we demonstrated that chimpanzees can recognize objects across dimensions using object form and/or features.

A comparison of the results of Phase 3 (Local Feature discrimination) to Phase 4 (Relative Position discrimination) reveals the chimpanzees' ability to recognize objects across dimensions using only their relative positions in space. The methods of Phase 4 removed both differential local feature and global form properties such that discrimination of the location of the hidden marshmallow could only be made based on the position of the two identical containers in relation to each other. Therefore, the success of all subjects in Phase 4 indicates that they can use relative position among objects in a 2D display to guide their 3D search.

In Phase 5 (Within Object discrimination: 2×1), the two identical containers of Phase 4 were combined to form a single object with two compartments along a horizontal plane. Discrimination of the location of the hidden marshmallow required identifying the correct location within the two-compartment object across dimensions. All chimpanzees demonstrated the ability to make cross-dimensional discriminations based on relative position within a single object.

Phases 6 through 8 (Within Object discrimination: 3×1, 2×2, 3×2) manipulated the complexity of the internal structure of the object and thus the amount of information that had to be obtained from the 2D image. Phase 6 (3×1) required subjects to make a 'left/middle/right' discrimination within the object across dimensions without the use of distinctive local features. Three of four subjects attained criterion on this phase. In Phase 7 (2×2), these three subjects were tested on their ability to discriminate the location of the hidden marshmallow in a more complex object again without the use of local features. Subjects had to discriminate between four featurally identical compartments across dimensions. All three subjects that attempted this phase attained criterion performance. In the final phase of testing (3×2), the complexity of the object was again increased and local features remained omitted. Therefore, subjects had to discriminate between six featurally identical compartments across dimensions to correctly locate the hidden marshmallow. One of three subjects attained criterion performance in this phase.

This disparity in performance among the chimpanzee subjects will be an important launching point for future research. One must consider what experiences may promote the use of more advanced stages of the proposed complex equivalence submode. Here, 3D control trials were used to rule out memory and/or general attentional failures leading to incorrect performance. Thus, failures are likely due to reinforcement contingencies, training, decreased attention to 2D environments and/or frustration effects. Further research is encouraged to tease apart the impact of these factors on performance and to present a clearer picture as to the microdevelopment and limitations of complex equivalence processing in nonhuman subjects.

An analysis of the distribution of errors for the three subjects that were presented with Phase 8 revealed significant differences between subjects that did and did not attain criterion performance in this phase. The subject that achieved criterion performance on this final phase (Lana) made eight errors, and these were not distributed between the 'Horizontal' and 'Vertical' type differently than expected by chance. Two of her errors (25%) were classified as 'Previous Location' and therefore may have been due to proactive interference effects (i.e. the location of marshmallow on previous trials interfered with the discrimination of its location on future trials). Errors of this type comprised 17% of total errors made by Sherman and Panzee (the subjects who did not attain criterion performance in this phase). Further, the distribution of Sherman's and Panzee's 'Horizontal' and 'Vertical' errors deviated significantly from that expected by chance; 'Horizontal' errors were more frequent than expected. Therefore, these subjects had greater difficulty discriminating between locations on the same horizontal row.

Converging evidence of a horizontal disadvantage in chimpanzees comes from work presented by Poti (2005) concerning spontaneous spatial constructions with objects in chimpanzees. Poti (2005) found that chimpanzees demonstrated the lowest level of sophistication of object combinations along the horizontal plane. Further evidence of a horizontal disadvantage to skill acquisition has been noted informally in the use of computerized testing systems with another species of nonhuman primate. Capuchin monkeys first learn the relationship between joystick movement and subsequent cursor displacement along the vertical axis, later followed by mastery of the task along the horizontal axis (personal observation). Object complexity along the horizontal plane should therefore be carefully considered prior to conducting cross-dimensional research with nonhuman primates.

Experiment 2

Method

Participants

Study participants were recruited in Athens, Georgia via parent email listserves provided by local daycares and

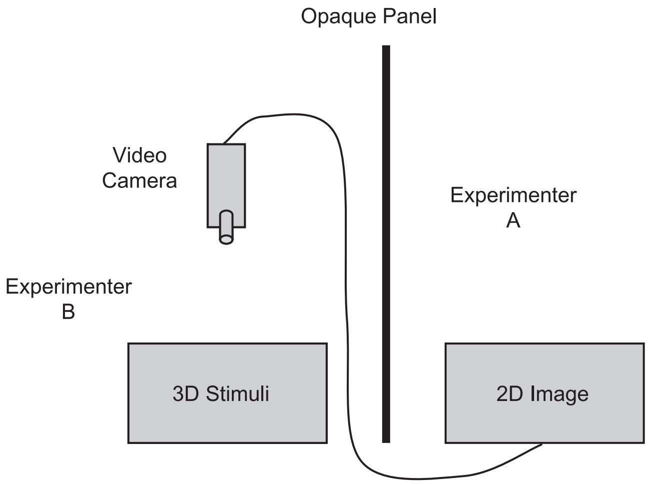

Figure 3 *Testing setup for toddlers.*

notices in family housing dormitories at the University of Georgia and in Orlando, Florida via community listserves and newspaper advertisements. Participants were assigned to one of two study groups, 3-year-olds and 4-year-olds. The 3-year-old group was composed of six males and six females each within 30 days of their third birthday. The 4-year-old group was made up of three males and nine females each within 30 days of their fourth birthday. Participants were predominantly Caucasian and from middle-class families. Parents completed a brief questionnaire to ensure that all participants had normal or corrected to normal vision and previous experience viewing televised images.

Materials

Testing was conducted in 2 m^2 of open floor space of the child's home. Identical experimental stimuli from Experiment 1 were used in this experiment with only slight modifications to the procedures. A cardboard partition (91.40 cm × 121.90 cm) was used to block the child's view of the 3D hiding events. Live images were captured using a Panasonic Hi-8 Camera and relayed to a 13″ Durabrand Color Television (see Figure 3).

Procedures

The participant was seated on the floor in front of the television monitor with experimenter 'A' (see Figure 3). Experimenter 'B' was positioned with the 3D experimental stimuli on the opposite side of the cardboard partition out of view of the child and experimenter 'A'. The video camera was positioned in front of the 3D experimental stimuli and sent a live image to the television screen. On each trial, experimenter 'B' would hide a small stuffed toy in the test object such that it was completely hidden from view as in Experiment 1. The child was instructed to view the actions of experimenter 'B' on the television and then Experimenter 'A' asked the child move around the opaque partition to side with the test objects. Experimenter 'B' called the child over and then asked the child to find the hidden toy.

All phases of testing were carried out using the same 3D stimuli and methods as in Experiment 1. To avoid biasing the choices of the participants, experimenter 'A' did not move to the test objects with the participant, and experimenter 'B' looked straight at the ground while the selection of the location was made. If the child found the hidden toy on the first choice he/she was allowed to carry the toy back to the television side of the partition and play with it prior to the initiation of the next trial.

Three-year-olds began the testing series with Phase 1 (Color and Form MTS) (see Figure 2a). To avoid redundancy, the 4-year-olds began the testing series with Phase 3 (Local Feature Discrimination) (see Figure 2c). To reduce the time demand on each participant's family as well as decrease testing fatigue, criterion performance for successful completion of a testing phase was altered from that used with chimpanzee subjects in Experiment 1. Criterion for human participants was defined as the production of the correct response on four consecutive trials within a single testing session. Upon successful completion of each testing phase the child was allowed to select a sticker from a sticker book. During each session a maximum of 10 trials per phase were administered. If criterion on a phase was not attained, this phase was administered during the next testing session. If the participant did not attain criterion on a phase after two test sessions (i.e. maximum of 20 trials) or no longer wished to participate, cross-dimensional testing was halted.

As in Experiment 1, 3D control trials were conducted to ensure that performance deficits in the cross-dimensional task were not due to memory or attentional failures. During these trials, the participant was presented with the phase of the task he/she could not complete in the cross-dimensional form. Participants directly viewed the experimenter hide the toy on one side of the partition (3D). The experimenter then moved the items to the opposite side of the partition and the participant would follow and make his/her choice of the location of the hidden item (3D). All participants attained criterion performance of four consecutive correct choices in the 3D control tests within five trials. At the completion of testing, each participant received a coloring book and certificate of participation.

Results

Phase 1: Color and Form MTS

All 12 participants in the 3-year-old age group attained criterion performance on the cross-dimensional matching-to-sample procedure in which the correct stimulus could be identified by properties of color and form. The mean number of trials needed to attain this criterion was 5.00 (SD = 1.95) (see Table 4).

Table 4 *Mean number of trials presented to attain criterion performance* in toddler groups*

Task	Age groups	
	3 yr olds	4 yr olds
1 Color and Form MTS	5.00 (*SD* = 1.95)	X
2 Form MTS	4.70 (*SD* = 1.06)	X
3 Local Feature discrimination	6.11 (*SD* = 3.89)	4.08 (*SD* = 0.29)
4 Relative Position discrimination	10.00 (*SD* = 0.00)	4.42 (*SD* = 1.00)
5 Within Object discrimination (2 × 1)	CNA	6.42 (*SD* = 4.23)
6 Within Object discrimination (3 × 1)	X	6.50 (*SD* = 3.09)
7 Within Object discrimination (2 × 2)	X	5.50 (*SD* = 2.32)
8 Within Object discrimination (3 × 2)	X	6.75 (*SD* = 4.22)

CNA = Criterion Not Achieved.
X = Phase Not Attempted.
* Criterion Performance = Correct on 4 consecutive trials in single session.

Phase 2: Form MTS

Ten of the 12 participants in the 3-year-old age group attained criterion performance on the cross-dimensional match-to-sample procedure in which the correct stimulus could be identified by form properties alone. The mean number of trials needed to attain this criterion was 4.70 (*SD* = 1.06) (see Table 4).

Phase 3: Local Feature discrimination

This cross-dimensional discrimination procedure in which the correct stimulus could be identified by a distinctive local feature (i.e. circle or triangle) or its relative position (i.e. right or left) served as the first phase of testing for the 4-year-old group and the third testing phase for the 3-year-old group. Of the 10 participants from the 3-year-old group that attempted this task, nine attained criterion performance. The mean number of trials needed to attain criterion performance was 6.11 (*SD* = 3.89) (see Table 4). All 12 of the participants from the 4-year-old group attained criterion performance. The mean number of trials to criterion among the 4-year-old group was 4.08 (*SD* = 0.29) (see Table 4).

Phase 4: Relative Position discrimination

Nine of the 3-year-old participants and all 12 of the 4-year-old participants attempted this cross-dimensional discrimination procedure in which distinctive local features were removed and the correct object could be identified by its relative position only. Only one participant from the 3-year-old age group attained criterion performance in the cross-dimensional version of the task and did so in 10 trials (see Table 4). 3D control trials were conducted with the other eight 3-year-old participants and all were able to attain criterion performance in the first five trials presented. Subsequent cross-dimensional testing was therefore halted with these eight 3-year-old participants. Among the 4-year-old participants, all 12 reached criterion on the cross-dimensional form of this phase. The mean number of trials to criterion among the 4-year-old group was 4.42 (*SD* = 1.00) (see Table 4).

Phase 5: Within Object discrimination – (2 × 1)

The one 3-year-old participant that attempted this testing phase did not attain criterion performance in the cross-dimensional form of the task, but did attain criterion performance in the 3D control phase in the first four trials presented. Subsequent cross-dimensional testing was therefore halted with this final 3-year-old participant. All 12 of the 4-year-old participants attained criterion performance in this phase. The mean number of trials to criterion among the 4-year-old group was 6.42 (*SD* = 4.23) (see Table 4).

Phases 6 to 8: Within Object discrimination – (3 × 1) to (3 × 2)

All 12 of the 4-year-old participants attained criterion performance on Phases 6 through 8 of the cross-dimensional search task. The mean number of trials to criterion and standard deviation by phase is reported in Table 4. An analysis of errors in Phase 8 was not conducted as in Experiment 1 due to a lack of errors made by the participants.

Discussion

Significant developmental differences in the ability to gain knowledge of an object's compositional structure from its 2D image were evident between the 3- and 4-year-old age groups. These findings demonstrate the importance of considering participant characteristics prior to conducting tasks using 2D stimuli. Previous findings suggest that young children comprehend the function and content of photographs and video displays at 30 months of age (DeLoache & Burns, 1994; Troseth & DeLoache, 1998). Findings from this series of tests imply that it is not until 4 years of age (i.e. 48 months) that children reliably gain knowledge of object structure from 2D depictions (i.e. operate in the proposed complex equivalence submode of 2D processing).

All 12 participants in the 4-year-old age group were presented with and attained criterion on Phase 3 (Local Feature Discrimination) through Phase 8 (Within Object Discrimination: 3×2). These results demonstrate that 4-year-olds can view an object's 2D image and gain knowledge of that object to guide their actions in the 3D world without relying on distinctive local features. Additionally, the attainment of criterion performance on later phases of the testing series by all participants in this age group demonstrates that the ability to discriminate between featurally identical locations within a single object from information gained from its 2D image exists even when the object becomes increasingly complex structurally. Therefore, it is clear from these findings that 4-year-old children process 2D images using the proposed complex equivalence submode.

The results from the 3-year-old group are in sharp contrast with those of the 4-year-old group. All 3-year-olds attained criterion on Phase 1 (Color and Form MTS). Thus, at 3 years of age children are able to recognize an object from a 2D depiction using characteristics of color or form. Further, 10 of 12 participants presented with Phase 2 (Form MTS) attained criterion performance. These participants were therefore able to recognize an object across dimensions without the use of the characteristic of color. In Phase 3 (Local Feature Discrimination), nine of 10 3-year-old participants presented with the task attained criterion. Thus, these participants could discriminate between two globally identical containers using a single distinctive local feature (circle vs. triangle) or the relative position of the objects in space. When moved on to Phase 4 (Relative Position Discrimination), only one of the nine 3-year-olds achieved criterion performance. This suggests that most 3-year-old participants were relying on the distinctive local feature to identify objects across dimensions in Phase 3. The one participant that was able to attain criterion performance on Phase 4 when the distinctive local features were removed could not do so in Phase 5 when the two objects were combined to form a single unit. These results lead us to conclude that 3-year-olds are generally operating in the proposed featural equivalence submode of processing 2D depictions of 3D objects.

The developmental difference in performance leads one to consider what ontological experiences are taking place between the ages of 3 and 4 years that promote the use of the complex equivalence submode of processing 2D stimuli. As with the chimpanzee subjects, the 3D control trials presented in this study ruled out memory and general attentional deficits as the cause of failures in testing phases requiring processing of 2D stimuli in the complex equivalence submode by 3-year-old participants. We encourage future research to address the identification of experiences that promote relational knowledge such as that required to discriminate between featurally identical locations within an object across dimensions.

The findings of this study have important implications for research conducted using 2D stimuli in lieu of 3D objects, particularly in the area of object recognition. We advise caution when interpreting findings collected from such experiments with children less than 4 years of age, in that they appear to rely on distinctive local features to solve cross-dimensional tasks.

General discussion

The findings of these two experiments help answer questions of internal validity that arise within the object recognition literature when 2D depictions are presented as stimuli in the place of 3D objects. From previous research on the perception of individual objects, it was unclear whether humans and nonhuman primate subjects completed such tasks by gaining knowledge of object structure from their 2D images or if they followed a response rule guided by local features across dimensions. In some cases it seemed as if subjects may have confused 2D and 3D items, demonstrated by trying to interact with an item depicted in two dimensions. That is, they could not distinguish between the two forms of presentation.

The current studies build upon a useful theoretical framework proposed by Fagot *et al.* (2000). These researchers proposed three modes of processing the relationship between 3D objects and their 2D images. Subjects may process these two forms independently, that is they may not perceive the representational relationship between the two. Subjects may confuse the 2D image and the 3D object in that they do not perceive there to be a difference between the two forms, demonstrated by attempts to interact with 2D depictions of objects in the same way as the 3D objects themselves. Finally, subjects may equate their perception from the 2D image with that of the 3D object. That is, they comprehend the symbolic relationship that exists between the 2D image and 3D object. We sought to demonstrate that the equivalence mode is composed of two submodes; featural and complex. In the featural equivalence submode, cross-dimensional tasks can be solved by following a response rule focused on distinctive local features. In the complex equivalence submode, knowledge of the object's global form and constituent structure is gained from the 2D depiction. This submode does not rely solely on local features and can be used to discriminate between relative positions within and between objects.

The results of the current studies demonstrate that the proposed submodes capture important individual differences in both humans and chimpanzees. Data from chimpanzee subjects revealed that all subjects were able to discriminate between identical locations within an object across dimensions without the use of distinctive local features. Additionally, no behaviors were observed that would lead one to suspect that the subject confused the 2D image and 3D object (i.e. reaching for the screen). These findings rule out chimpanzees' processing

Table 5 *Number of subjects that achieved criterion performance/number of subjects presented phase*

	Experimental groups		
Task	Human – 3 yr	Human – 4 yr	Chimpanzees
1 Color and Form MTS	12/12	X	4/4
2 Form MTS	10/12	X	4/4
3 Local Feature discrimination	9/10	12/12	4/4
4 Relative Position discrimination	1/9	12/12	4/4
5 Within Object discrimination (2 × 1)	0/1	12/12	4/4
6 Within Object discrimination (3 × 1)	X	12/12	3/4
7 Within Object discrimination (2 × 2)	X	12/12	3/3
8 Within Object discrimination (3 × 2)	X	12/12	1/3

of the relationship between an object and its 2D depiction using the independence or confusion modes. Since the chimpanzees were able to attain criterion performance without the use of distinctive local features, the featural equivalence submode of processing can also be excluded. The chimpanzees in this study therefore appear to have been operating using the complex equivalence submode to varying degrees. One subject was able to use structural knowledge of the most complex object presented (3 × 2) from its 2D depiction. Two chimpanzees attained criterion performance with the slightly less complex (2 × 2) object, and the final subject was able to make left/right discriminations within a single (2 × 1) object without the use of distinctive local features.

In contrast to the chimpanzees, 3-year-old human participants generally operated in the featural equivalence submode. That is, they demonstrated the ability to recognize globally distinct objects across dimensions, thus ruling out their use of the independence mode. No behaviors were observed from 3-year-old participants indicating that they were confusing the 2D image with the 3D objects (i.e. reaching for the toy on the TV screen), therefore ruling out the use of the confusion mode. While nine of 10 participants that were presented with Phase 3 (Local Feature Discrimination) were able to attain criterion, only one did so when the distinctive local features were removed in Phase 4 (Relative Position Discrimination). This participant was then unable to complete the task successfully when the two distinct hiding locations were combined to form a single object in Phase 5. These results suggest that 3-year-old participants generally match 2D images with 3D objects using the featural equivalence submode.

The performance of 4-year-old participants was in sharp contrast to that of the 3-year-old group. All 4-year-olds attained criterion performance on Phase 8 (Within Object Discrimination: 3 × 2), demonstrating their ability to gain knowledge of the internal structure of an object from its 2D image without the use of local features. Thus, the 4-year-old group clearly operated cross-dimensionally using the complex equivalence submode. There is thus a significant developmental difference in the ability to use 2D images to act in the 3D world between the ages of 3 (featural equivalence) and 4 (complex equivalence). This finding suggests caution when interpreting results of cross-dimensional research conducted with children younger than 4 years of age.

The performance of chimpanzees more closely resembles that of 4-year-olds (complex equivalence) than that of 3-year-old participants (featural equivalence) (see Table 5). While one chimpanzee used an advanced level of the complex equivalence submode to solve the final task presented here (3 × 2), lower levels of this mode of processing were found in the three other subjects. Thus, the chimpanzees demonstrated the capacity to utilize the complex equivalence submode to varying degrees in these object perception tasks. However, one should not assume that the most sophisticated form of this process is being utilized by all chimpanzee subjects in all situations.

The disparity that exists in the results of individual chimpanzee subjects and between the 3- and 4-year-old toddler groups leads to a number of questions for future research. First, we must address the types of experiences that afford learning to process 2D stimuli in the complex equivalence submode. Second, we should determine whether there are attentional demands that are specific to the 2D environment. Further, while a vast literature exists addressing the development of visuomotor skills in humans and nonhumans (e.g. locomotion, spatial and object knowledge, goal directed behavior), the theoretical frameworks used to study them are difficult to apply to 2D environments. Theories such as the ecological approach to perception proposed by Gibson (1986) rely on perceptual feedback from all sensory modalities for skill learning. In the 2D environment, perceptual feedback outside of the visual domain is absent and the visual cues are reduced. Thus, these environments afford quite different perceptual experiences. Therefore, the application of theories describing skill development in 3D does not accurately represent perceptual experiences within the 2D environment. Finally, the use of the 2D environment both in scientific investigation (2D stimulus displays) and everyday life (reading, television, video games, computers) is becoming increasingly prevalent. Thus, we encourage further discussion of these topics to strengthen links between developmental frameworks and studies of the perception of 2D stimuli and environments.

It is important to note that the 2D presentation used in this study was a live televised image. Because this televised image displayed motion, it provided additional

3D cues that would not be present in a still photograph. Future research using this paradigm in conjunction with photographs will be necessary to determine the impact of 3D cues on performance. Further, the results of this study inform us of the subject's ability to use information from a 2D display to act in the 3D world along a single plane. That is, all discriminations in location were made within the same depth plane. The findings of Brown (1969) suggest that the perception of depth from 2D images may not develop until 6 years of age. The testing paradigm presented here could be altered to examine this ability developmentally and across species.

Acknowledgements

This work was support by NIH grants HD-38051, MH-58855, HD-056352, a NIH pre-doctoral NRSA MH-070090, Sigma Xi, and a PEO Scholar's Award. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH. We would like to thank John Kelley, Sydney Felker, Dr Hika Kuroshima, and Dr Erin Riley for their assistance in data collection. Thanks to Dr James Brown, Dr Michael Beran, and anonymous reviewers for their theoretical insights and discussion of this research.

References

Bovet, D., & Vauclair, J. (1998). Functional categorization of objects and of their pictures in baboons (*Papio anubis*). *Learning and Motivation*, **29**, 309–322.

Bovet, D., & Vauclair, J. (2000). Picture recognition in animals and humans. *Behavioural Brain Research*, **109**, 143–165.

Boysen, S.T., & Berntson, G.G. (1986). Cardiac correlates of individual recognition in the chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, **100**, 321–324.

Brakke, K.E., & Savage-Rumbaugh, E.S. (1995). The development of language skills in bonobo and chimpanzee – I. Comprehension. *Language and Communication*, **15**, 121–148.

Brakke, K.E., & Savage-Rumbaugh, E.S. (1996). The development of language skills in Pan – II. Production. *Language and Communication*, **16**, 361–380.

Brown, L.B. (1969). The 3D reconstruction of a 2D visual display. *The Journal of Genetic Psychology*, **115**, 257–262.

Davenport, R.K., & Rogers, C.M. (1971). Perception of photographs by apes. *Behaviour*, **39**, 318–320.

DeLoache, J.S. (2000). Dual representation and young children's use of scale models. *Child Development*, **71**, 329–338.

DeLoache, J.S., & Burns, N.M. (1994). Early understanding of the representational function of pictures. *Cognition*, **52**, 83–110.

DeLoache, J.S., Pierroutsakos, S.L., Uttal, D.H., Rosengren, K.S., & Gottlieb, A. (1998). Grasping the nature of pictures. *Psychological Science*, **9**, 205–210.

Deregowski, J.B., Muldrow, E.S., & Muldrow, W.F. (1972). Pictorial recognition in a remote Ethiopian population. *Perception*, **1**, 417–425.

Fagot, J., Martin-Malivel, J., & Dépy, D. (2000). Equivalence between objects and pictures. In J. Fagot (Ed.), *Picture perception in animals* (pp. 295–320). Hove, East Sussex: Psychology Press.

Gibson, J.J. (1986). *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Herzog, M., & Hopf, S. (1986). Recognition of visual pattern components in squirrel monkeys. *European Archives of Psychiatry and Neurological Sciences*, **236**, 10–16.

Hopkins, W.D., Washburn, D.A., & Hyatt, C.W. (1996). Video-task acquisition by rhesus monkeys (*Macaca mulatta*) and chimpanzees (*Pan troglodytes*): a comparative analysis. *Primates*, **37**, 197–206.

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., & Mukobi, K.L. (1999). Scale-model comprehension by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, **113**, 396–402.

Kyes, R.C., Mayer, K., & Bunnell, B.N. (1992). Perception of stimuli presented as photographic slides in cynomolgus macaques (*Macaca fascicularis*). *Primates*, **33**, 407–412.

Malone, D.R., Tolan, J.C., & Rogers, C.M. (1980). Cross-modal matching of objects and photographs in the monkey. *Neuropsychology*, **18**, 693–697.

Martin-Malivel, J. (1998). Existe-t-il une équivalence entre l'objet et sa représentation photographique ćhez le babouin (*Papio papio*)? *Primatologie*, **1**, 249–268.

Martin-Malivel, J., & Fagot, J. (2001). Cross-modal integration and conceptual categorization in baboons. *Behavioural Brain Research*, **122**, 209–213.

Menzel, C.R. (2001). A chimpanzee's learning of real-world events through second-order signals. *American Journal of Primatology*, **54** (Suppl. 1), 100.

Menzel, C.R. (2005). Progress in the study of chimpanzee recall and episodic memory. In H. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 188–224). New York: Oxford University Press.

Menzel, E.W., Premack, D., & Woodruff, G. (1978). Map reading by chimpanzees. *Folia Primatologica*, **29**, 241–249.

Menzel, E.W., Savage-Rumbaugh, E.S., & Lawson, J. (1985). Chimpanzee (*Pan troglodytes*) spatial problem solving with the use of mirrors and televised equivalents of mirrors. *Journal of Comparative Psychology*, **99**, 211–217.

Miller, R.J. (1973). Cross-cultural research in the perception of pictorial materials. *Psychological Bulletin*, **80**, 135–150.

Pierroutsakos, S.L., & DeLoache, J.S. (2003). Infants' manual exploration of pictorial objects varying in realism. *Infancy*, **4**, 141–156.

Plimpton, E.H., Swartz, K.B., & Rosenblum, L.A. (1981). Responses of juvenile bonnet macaques to social stimuli presented through color videotapes. *Developmental Psychobiology*, **14**, 109–115.

Poss, S.R. & Rochart, P. (2003). Referential understanding of videos in chimpanzees (*Pan troglodytes*), orangutans (*Pongo pygmaeus*), and children (*Homo sapiens*). *Journal of Comparative Psychology*, **117**, 420–428.

Poti, P. (2005). Chimpanzees' constructional praxis (*Pan paniscus*, *P. troglodytes*). *Primates*, **46**, 103–113.

Premack, D., & Woodruff G. (1978). Chimpanzee problem-solving: a test for comprehension. *Science*, **202**, 532–535.

Reid, S.L., & Spetch, M.L. (1998). Perception of pictorial depth cues by pigeons. *Psychonomic Bulletin and Review*, **5**, 698–704.

Rilling, M., LaClaire, L., & Warner, M. (1993). A comparative, hierarchical theory for object recognition and action. In T.R. Zentall (Ed.), *Animal cognition: A tribute to Donald A. Riley* (pp. 313–333). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rose, S.A. (1977). Infants' transfer of response between 2D and 3D stimuli. *Child Development*, **48**, 1086–1091.

Rosenfeld, S.A., & van Hoesen, G.W. (1979). Face recognition in the rhesus monkey. *Neuropsychologicia*, **17**, 503–509.

Rumbaugh, D.M. (1977). *Language learning by a chimpanzee: The LANA project*. New York: Academic Press.

Rumbaugh, D.M., Washburn, D.A., Savage-Rumbaugh, E.S., & Hopkins, W.D. (1991). Language Research Center's computerized test system (LRC-CTS): video formatted tasks for comparative primate research. In A. Ehara, T. Kimura, O. Takenake, & M. Iwatoto (Eds.), *Primatology today* (pp. 325–328). Amsterdam: Elsevier Science.

Sackett, G.P. (1965). Response of rhesus monkeys to social stimulation by means of coloured slides. *Perceptual and Motor Skills*, **20**, 1027–1028.

Sackett, G.P. (1966). Monkeys reared in isolation with pictures as visual input: evidence for an innate releasing mechanism. *Science*, **154**, 1468–1472.

Savage-Rumbaugh, E.S. (1986). *Ape language: From conditioned response to symbol*. New York: Colombia University Press.

Savage-Rumbaugh, E.S., Rumbaugh, D.M., Smith, S.T., & Lawson, J. (1980). Reference: the linguistic essential. *Science*, **210**, 922–925.

Shyi, G.C., & Huang, S.T. (1995). Constructing 3D mental models from viewing 2D displays. *Chinese Journal of Psychology*, **37**, 101–122.

Slater, A., Rose, D., & Morison, V. (1984). New-born infants' perception of similarities and differences between two- and three-dimensional stimuli. *British Journal of Developmental Psychology*, **2**, 287–294.

Tanaka, M. (1996). Information integration about object–object relationships by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, **110**, 323–335.

Troseth, G.L., & DeLoache, J.S. (1998). Medium can obscure the message: young children's understanding of video. *Child Development*, **69**, 950–965.

Watanabe, S. (1997). Visual discrimination of real objects and pictures by pigeons. *Animal Learning and Behavior*, **25**, 185–192.

Winner, E., & Ettlinger, G. (1979). Do chimpanzees recognize photographs as representations of objects? *Neuropsychologia*, **17**, 413–420.

Zimmerman, R.R., & Hochberg, J. (1970). Responses of infant monkeys to pictorial representations of a learned visual discrimination. *Psychonomic Science*, **18**, 307–308.

Zimmerman, R.R., & Hochberg, J. (1971). The facilitation of picture discrimination after object discrimination learning in the neonatal monkey and vice versa. *Psychonomic Science*, **24**, 239–241.

Received: 4 May 2006
Accepted: 2 October 2007