Published in final edited form as:

Cognition. 2016 May ; 150: 53–67. doi:10.1016/j.cognition.2016.01.012.

Smoke and mirrors: Testing the scope of chimpanzees' appearance-reality understanding

Carla Krachun^a, Robert Lurz^b, Jamie L. Russell^{c,d}, and William D. Hopkins^{c,d}

Carla Krachun: ckrachun@grenfell.mun.ca; Robert Lurz: rlurz@brooklyn.cuny.edu; Jamie L. Russell: jlcox@emory.edu; William D. Hopkins: whopkins4@gsu.edu

^aGrenfell Campus, Memorial University of Newfoundland, 20 University Drive, Corner Brook, NL, Canada, A2H 5G4

^bBrooklyn College, CUNY, 2900 Bedford Avenue, Brooklyn, NY 11210

^cGeorgia State University

^dYerkes National Primate Research Center, 954 Gatewood Rd., Atlanta, GA 30329

Abstract

The ability to make appearance-reality (AR) discriminations is an important higher-order cognitive adaptation in humans but is still poorly understood in our closest primate relatives. Previous research showed that chimpanzees are capable of AR discrimination when choosing between food items that appear, due to the effects of distorting lenses, to be smaller or larger than they actually are (Krachun, Call & Tomasello, 2009). In the current study, we investigated the scope and flexibility of chimpanzees' AR discrimination abilities by presenting them with a wider range of illusory stimuli. In addition to using lenses to change the apparent size of food items (Experiment 1), we used a mirror to change the apparent number of items (Experiment 2), and tinted filters to change their apparent color (Experiment 3). In all three experiments, some chimpanzees were able to maximize their food rewards by making a choice based on the real properties of the stimuli in contrast to their manifest apparent properties. These results replicate the earlier findings for size illusions and extend them to additional situations involving illusory number and color. Control tests, together with findings from previous studies, ruled out lower-level explanations for the chimpanzees' performance. The findings thus support the hypothesis that chimpanzees are capable of making AR discriminations with a range of illusory stimuli.

Keywords

Appearance-reality discrimination; visual illusions; chimpanzees; nonhuman primates

---

Correspondence to: Carla Krachun, ckrachun@grenfell.mun.ca.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. INTRODUCTION

Every day, we confront a variety of situations in which things are not as they appear to be. These range from simple perceptual illusions, such as when a straight stick in water looks bent, to more complex social interactions, such as when a person who is lying appears to be telling the truth. These kinds of situations challenge us to draw upon an ability that begins to develop early in life and becomes a crucial component of the human cognitive skill set: the ability to make appearance-reality (AR) discriminations (Flavell, Flavell & Green, 1983). Understanding this distinction allows us to behave more adaptively in everyday interactions with our physical and social worlds—it helps us to avoid being misled by deceptive objects or people, and it gives us a tool for manipulating others to suit our own ends. Indeed, AR discrimination forms the foundation of all science and objective pursuits of truth, for without it we could make no sense of the distinction between fact and opinion (Braine & Shanks, 1965; Davidson, 2004; Nagel, 2012; Strawson, 1959). In addition, it may be the very basis on which we come to know about the existence of different subjective points of views (Flavell, Green & Flavell, 1986; Gopnik & Astington 1988; Moll, Meltzoff, Merzsch & Tomasello, 2012).

The human literature on AR discrimination reflects the enormous significance of this cognitive skill. A considerable amount of research has focused on determining when and how it develops, and its relationship to other important cognitive accomplishments such as perspective-taking and false-belief understanding (for reviews see Flavell, 1986, 1993). Many studies have used paradigms that test children's developing ability to discriminate the real from the apparent nature of viewed objects. In 'identity illusions' an object appears to be one thing but is truly another; for example, a sponge looks like a rock, or an eraser looks like a piece of chocolate (e.g., Flavell et al., 1983). In 'property illusions' the physical appearance of an object is transformed by illusion-producing stimuli such as size-altering lenses or color filters (e.g., Braine & Shanks, 1965; Flavell et al., 1986). The standard paradigm in early studies involved presenting children with the object, allowing them to interact with it to discover its illusory nature, and then asking the child to verbally report on its apparent and true identity or properties.

The early work by Flavell and others (Gopnik & Astington, 1988; Taylor & Hort, 1990) pointed to a developmental shift toward reliably passing AR tests at 4–5 years of age. Researchers concluded that this shift reflected a major conceptual change whereby children begin to understand that the same object, event, or situation can be construed in two different ways simultaneously, either by the same person or by two different people (Flavell, 1993; Perner, 1991). This conceptual change has been hypothesized to be responsible for a similar developmental shift that occurs at the same age in other tests involving dual representation, including tests of false-belief understanding and level-2 perspective taking (Andrews, Halford, Bunch, Bowden & Jones, 2003; Flavell, 1993; Gopnik & Astington, 1988; Perner, 1991; Wellman 1990). However, others argue that younger children's difficulties with standard AR tests have less to do with dual-representation limitations than with the high linguistic or memory demands of the tasks (Deák, 2006; Hansen & Markman, 2005; Rice, Koinis, Sullivan, Tager-Flusberg & Winner, 1997; Sapp, Lee & Muir, 2000).

In contrast to the substantial human developmental research in this area, there is still very little known about the existence of AR discrimination in other animals, even in our closest primate relatives. Yet, other species no doubt encounter perceptual ambiguities in their environments, and there are reports from both observational studies (Byrne & Whiten, 1990; Hirata & Matsuzawa, 2001; Menzel, 1974; Mitchell & Thompson, 1986; Whiten & Byrne, 1988) and experimental studies (Hare, Call & Tomasello, 2006; Woodruff & Premack, 1979) of chimpanzees apparently trying to misrepresent themselves in order to deceive others. The capacity to respond adaptively in these sorts of illusory situations may have conferred a selective advantage in our early primate ancestors, forming the evolutionary basis for our own AR discrimination skills. There are further reasons to think that chimpanzees may be capable of AR discrimination. Chimpanzees are known to use mirrors to inspect parts of their own bodies (Gallup, 1970; Povinelli, Rulf, Landau & Bierschwale, 1993) and to guide their limbs towards hidden objects (Menzel, Savage-Rumbaugh & Lawson, 1985), demonstrating a recognition that what may appear to be another chimpanzee is actually just a reflection of themselves. Chimpanzees have also been found to be capable of using a scale model of their enclosure as a source of information about where food is hidden in the enclosure, an ability which arguably involves a form of dual representation (Kuhlmeier, Boysen & Mukobi, 1999). A number of chimpanzees (and other great apes) have also succeeded in Piagetian-style conservation tasks in which, for example, a fixed amount of liquid may appear to increase or decrease when it is transferred into another container (Muncer, 1983; Suda & Call, 2004; Woodruff, Premack & Kennel, 1978). That these apes were not fooled by the transformation suggested they could be capable of AR discrimination.

To date, there have been two published studies explicitly designed to test for AR discrimination in apes. In the first study, carried out with 14 chimpanzees at the Leipzig Zoo in Germany (Krachun et al., 2009), distorting lenses were used to test apes' ability to discriminate between the real and apparent size of grapes, a highly valued food item for chimpanzees. The animals were first given a simple test (called the 'basic test') in which they watched an experimenter place a large grape behind a minimizing lens so that it looked small, and a small grape behind a magnifying lens so that it looked large. Chimpanzees were then allowed to choose between the two grapes. Eight of the 14 chimpanzees chose the apparently small (but truly large) grape significantly more often than chance across two 12-trial sessions, suggesting they recognized that the appearance of the grapes was misleading. However, the researchers recognized that the animals could have solved the task simply by visually tracking the large grape throughout the procedure. To rule this out, successful chimpanzees were given a 'tracking control test,' in which a procedural modification made it impossible for them to engage in visual tracking. Five of the eight chimpanzees also passed that test within just 12 trials. Furthermore, none of the chimpanzees who passed the AR test were capable of passing either of two similarly structured reverse contingency control tests in which they had to pick a small grape to obtain a large one. Although the reverse contingency tests were designed to be as similar as possible to the AR test (but without the use of distorting lenses), chimpanzees' performance in these tests was well below chance and significantly worse than their performance in the corresponding sessions of the AR test (Krachun et al., 2009). These results agreed with findings from numerous

other studies (Boysen & Berntson, 1995; Boysen, Berntson, Hannan, & Cacioppo, 1996; Boysen, Berntson & Mukobi, 2001; Boysen, Mukobi, & Berntson, 1999; Uher & Call, 2008; Vlamings, Uher & Call, 2006), which have consistently shown that chimpanzees require a great deal of training (typically hundreds of trials) to master reverse contingency tests. Thus, it is highly unlikely that the chimpanzees in Krachun et al. (2009) solved the AR test by quickly learning a reverse contingency rule.

Recently, Karg, Schmelz, Call and Tomasello (2014) suggested a new way of testing apes' ability to discriminate between the real and apparent size of food items, and they administered their test to 29 great apes (including 13 chimpanzees) at the Leipzig Zoo. Their approach was to reverse the apparent relative size difference between two pretzel sticks by occluding a portion of each pretzel with a board, such that the longer pretzel had only a small section protruding from under the board whereas the shorter pretzel had a larger section showing. All apes received three conditions in counterbalanced order. In the 'reality-appearance' (RA) condition, apes were first shown the true sizes of the pretzel sticks by displaying them unoccluded on the table top; the board was then positioned over each pretzel, and the apes were allowed to choose. In the 'appearance-reality-appearance' (ARA) condition, apes first saw the partially occluded pretzels; the board was then removed to reveal the pretzels' true sizes and replaced again before allowing the apes to choose. In a control condition, apes were shown only the partially occluded pretzels before being allowed to choose. As a group, the apes performed very well in the RA and ARA conditions, choosing the truly longer pretzel in about 85% of trials, and there were no significant differences across species. On an individual level, over 80% of apes chose the truly longer pretzel significantly more often than chance over two 12-trial sessions in both experimental conditions (Karg, personal communication). Karg et al. (2014) attribute their higher success rate, relative to Krachun et al.'s (2009), to the greater ecological validity afforded by using partial occlusion to manipulate the apparent size of food items (since apes often encounter partially occluded objects in their natural environments).

However, since Karg and colleagues did not run a tracking control test, it is possible the apes' high success rate was due to their visually tracking the location of the larger pretzel stick, rather than their making an AR discrimination. To succeed in both the RA and ARA conditions, the apes need only have noted where the large pretzel was when it was not occluded and then chosen the pretzel in that location later. To be sure that the apes were not relying on such visual tracking to pass the test, the tracking explanation would need to be ruled out.

In summary, while much is known about AR discrimination in human children, very little is known about chimpanzees' capacity to distinguish real from apparent. The two published studies in this area are limited in a number of ways. In Krachun et al. (2009), just over a third of the chimpanzees tested passed both the basic test and the tracking control tests. Success rates were higher in Karg et al. (2014), but their results are difficult to interpret without a tracking control test. Further, both studies tested the same population of apes (at the Leipzig Zoo), and thus they cannot tell us the extent to which the abilities observed might also be present in other populations. Finally, both studies only investigated apes' ability to discriminate between the real and apparent *size* of items. This means that we

cannot know whether the AR discrimination skills that chimpanzees have thus far demonstrated are restricted to this particular property of items (perhaps because they are more likely to encounter size illusions in nature) or are more general in scope.

In the current study, our first step was to search for evidence of AR discrimination in a new population of chimpanzees, using a procedure based closely on the procedure used by Krachun et al. (2009). Beyond just replicating the 2009 results, however, our goal was to determine whether the success rate could be improved by giving the animals more of an opportunity to learn about the illusory properties of the stimuli. Given that the individuals tested by Krachun et al. (2009) had never before encountered distorting lenses, the criteria for passing used in that study may have been overly stringent. The apes were allowed just two sessions (12 trials each) to pass the basic test, two sessions (6 trials each) to pass the seen tracking test, and two sessions (6 trials each) to pass the unseen tracking test. We predicted that increasing the animals' exposure to the stimuli by increasing the number of experimental sessions and trials would better allow them to demonstrate any AR discrimination abilities they do possess.

Our second, and main, goal for the current study was to determine the scope and flexibility of chimpanzees' AR discrimination abilities by testing them with a wider range of illusory stimuli. As noted previously, children have been tested with a variety of stimuli, from distorting lenses to color filters to objects with illusory or pretend identities. This variety has allowed researchers to determine whether the developmental picture looks similar across different types of tests (e.g., Flavell, Flavell & Green, 1987), and to pinpoint superfluous task demands that may influence performance (e.g., Friend & Davis, 1993). Expanding the range of stimuli in AR tests with chimpanzees will make similar comparisons possible, both within and across species.

With the above goals in mind, we ran three AR tests in three separate experiments. In the *Lens Test* (Experiment 1), we used magnifying and minimizing lenses to investigate chimpanzees' ability to recognize when an object looked smaller or larger than it actually was. In the *Mirror Test* (Experiment 2), we used a mirror to investigate chimpanzees' ability to recognize when the number of items in a group looked like more than it actually was. In the *Color Test* (Experiment 3), we used transparent color filters to investigate chimpanzees' ability to recognize when an object looked to be a different color than it actually was. Each experiment employed a simple choice paradigm fashioned after the procedure used by Krachun et al. (2009). Chimpanzees were given a choice between two items, one of which appeared to be the more desirable item but was in reality the less desirable one, and vice versa. Chimpanzees who could discriminate between the apparent and real properties of the objects would be successful in choosing the truly more desirable item. In all three experiments, chimpanzees were tested in a stepwise fashion, with only those chimpanzees who passed a given stage of testing advancing to the next stage. While the procedures varied somewhat across the three tests, each progressed in the same general fashion. First, preference tests were administered to confirm that the chimpanzees had any natural preferences required for participation. For example, in Experiments 1 and 2, the chimpanzees needed to exhibit a preference for larger amounts of food over smaller amounts. Any special pretest training that was required by the test was also given at this

time. Specifically, in Experiment 3, the chimpanzees needed to learn that a white box was always empty whereas a color (green or yellow) box always contained food.

Chimpanzees who passed the preference testing and training then received their first introduction to the illusion-producing stimuli (lenses/mirrors/color filters), and a *basic test* similar to Krachun et al.'s (2009) was run. The goal of this test was to determine whether the chimpanzees were capable of ignoring the misleading appearance of the items in their illusory contexts and choosing the truly more desirable item. Chimpanzees who passed the basic test were then given a *tracking control test* (consisting of a 'seen' component and an 'unseen' component) to rule out the possibility that they had learned to succeed by visually tracking the location of the more desirable item. Finally, chimpanzees who passed the tracking control were given the *avoid-stimulus test* (called the *avoid-lens*, *avoid-mirror*, or *avoid-filter test*, depending on the illusory stimuli used) to rule out the possibility that they had succeeded by learning to avoid the illusion-producing stimulus itself. With these two possible alternative strategies ruled out, we could be more confident that the chimpanzees were making an appearance-reality discrimination.

2. EXPERIMENT 1: LENS TEST

In this first experiment, we sought to replicate the results of Krachun et al. (2009) with a different population of chimpanzees, and to determine whether higher success rates could be achieved with greater opportunity to learn about the properties of the lenses.

2.1. Method

2.1.1. Subjects—Five chimpanzees housed socially at the Yerkes National Primate Research Center in Atlanta, Georgia participated in Experiment 1 (see Table 1). They were tested in the outdoor portions of their home enclosures; were fed a regular diet of fruits, vegetables, nuts, grains and primate chow; had water available *ad libitum*; and participated voluntarily in all sessions by approaching the experimenter when called. The chimpanzees had previously participated in a number of studies on handedness, tool-use, social cognition, and physical cognition (Hopkins, Adams & Weiss, 2013; Hopkins, Keebaugh, et al., 2014; Hopkins, Reamer, Mareno & Schapiro, 2015; Hopkins, Russell & Schaeffer, 2014; Russell, Lyn, Schaeffer & Hopkins, 2011), but they had no previous experience with AR discrimination tasks.

2.1.2. Design—The experiment included five tests presented in a fixed order: (1) the preference test with a maximum of two 12-trial sessions; (2) the basic test with a maximum of eight 12-trial sessions; (3) the seen tracking test with a maximum of four 12-trial sessions; (4) the unseen tracking test with a maximum of four 12-trial sessions; and (5) the avoid-lens test with a maximum of four 12-trial sessions. For each of the tests (excluding the preference test), each session began with 2–4 warmup preference trials (which followed the same procedure as in the initial preference testing) followed by 2 demonstration (*demo*) trials. These were followed immediately by the 12 test trials. In every test, the larger grape was on the left and right side an equal number of times, its position was randomized across trials, and it was never in the same position for more than two consecutive trials. Testing proceeded in a stepwise fashion, such that the chimpanzees needed to pass one test in order

to advance to the next test in the sequence. Passing a test was defined as choosing the larger grape significantly more often than chance within the maximum number of sessions allowed for that test (see results section for details on passing criteria).

2.1.3. Experimental stimuli—The illusory stimuli for this experiment consisted of two containers (16 X 16 X 13 cm) with distorting lenses ($d = 6$ cm) mounted into their faces (17 X 10.5 cm), similar to those used by Krachun et al. (2009). In the magnifying container, the lens caused a small grape (2 X 1.5 cm) placed behind it to appear large; in the minimizing container, the lens caused a large grape (2.75 X 1.75 cm) placed behind it to appear small (see Figure 1a). The noted sizes of the small and large grapes are approximate measures of length and width. We selected these sizes of grapes for two reasons: (1) preference testing confirmed that the chimpanzees could detect such a size difference and did prefer the larger grape; and (2) when placed behind the lenses, the truly smaller grape appeared the same size as the truly larger grape did in ‘real life’, and vice versa (see Figure 1a). Grapes were initially measured and selected using a ruler, and then used as exemplars against which additional grapes were visually compared and selected.

2.1.4. Procedure—All chimpanzees were tested individually. During testing, they sat facing an experimenter (E) across a table (height = 50 cm) positioned in front of their enclosure. The stimuli were displayed on the table (approximately 30 cm apart), and chimpanzees could make a choice by poking a finger through the wire mesh and pointing to one of the containers.

2.1.4.1. Preference test: First, we needed to confirm that the chimpanzees had a natural preference for larger grapes over smaller grapes. To test this, E presented them with a simple choice between two different-sized grapes displayed on a table top. All chimpanzees showed a statistically significant preference for the larger grape within two testing sessions and thus proceeded to the next stage of testing. As noted above under Design (section 2.1.2), two warmup preference trials were repeated at the start of each subsequent testing session. Besides serving as a warmup to the session, these trials allowed us to confirm chimpanzee motivation at that time. When chimpanzees chose the larger grape in both of these trials, we proceeded with testing. When they did not choose the larger grape in both trials, E administered two additional preference trials. In these cases, we proceeded with the test session only if the larger grape was chosen in at least three out of four trials. Otherwise, the session was postponed until later (this happened just twice).

2.1.4.2. Basic test: The basic test investigated whether chimpanzees were capable of ignoring the illusory appearance of the grapes and choosing the truly larger of the two. Each basic test session included the warmup preference trials (see above), two demo trials (described below), and 12 basic trials.

Demo trials: The chimpanzees tested had no prior experience with magnifying or minimizing lenses. Thus, to help them learn about the properties of the lenses, we followed the warmup trials in each basic test session with two trials in which we demonstrated the properties of the lenses for them. In these trials, E placed a small grape into the magnifying container and a large grape into the minimizing container, and then pushed the containers

forward so the chimpanzee could choose one. Once the chimpanzee made a choice, E demonstrated the properties of the lenses, as follows. She slowly flipped open the front of each container so that the chimpanzee could see the transformation in the appearance of the grapes as the lenses were removed. She did this at least twice for each container, starting with the chosen container. After E finished performing this demonstration, she gave the chimpanzee the chosen grape and discarded the other one. Two such demo trials were repeated at the start of every test session, to remind the chimpanzees about the properties of the lenses.

Basic trials: Immediately following the two demo trials, E ran 12 basic trials which were similar to the previous demo trials but did not include an extensive demonstration of the properties of the lenses. The procedure for these trials was as follows: E baited the containers as the chimpanzee watched, putting the small grape behind the magnifying lens and the large grape behind the minimizing lens. She then pushed the containers apart and forward on the table. Chimpanzees who were capable of choosing based on what they knew to be the true size of the grapes, rather than their apparent size, would be expected to choose the truly larger grape behind the minimizing lens. When the chimpanzee chose a container, E flipped open the front of both containers, gave the chimpanzee the chosen grape, and discarded the other grape. Chimpanzees who passed the basic test within eight sessions proceeded to the tracking control test.

2.1.4.3. Tracking control test: Chimpanzees who succeeded in the basic test demonstrated that they were capable of ignoring the misleading appearance of the grapes and choosing the one they knew to be truly larger. However, this alone does not mean that the chimpanzees were making an AR discrimination; they could have just visually tracked the larger grape as E baited the containers and positioned them on the table. To rule out this possibility, we had E stack the containers one on top of the other during the baiting process and then block the chimpanzee's view with an occluder before positioning the containers on the left and right sides of the table (see Figure 2). Thus, when the occluder was removed, the chimpanzee had to remember that the grape that looked larger was actually the smaller one, and vice versa. Remembering and applying such a rule could only be done if the chimpanzee was aware of the difference between how the grapes appeared and their real sizes—simply ignoring the misleading appearance of the grapes and tracking the large grape would not be enough. The tracking control test was run in two stages: the seen tracking test followed by the unseen tracking test. The purpose of the seen tracking test was to familiarize chimpanzees with the new stacking procedure while allowing them to observe the baiting and placement of the containers. It thus served as a transition to the truly diagnostic unseen tracking test, in which chimpanzees were prevented from visually tracking the food items by occluding their view during the placement of the containers.

Seen tracking test: The procedure for this test was very similar to the basic test procedure described above except that E stacked the containers in the middle of the table as she baited them. Thus, E baited one container in the middle of the table as the chimpanzee watched, then stacked the second container on top of the first and baited that one. She then unstacked the containers, positioned them on the left and right sides of the table, and pushed them

forward so that the chimpanzee could choose. The position of the container with the truly large grape inside was randomly determined, with the stipulation that it be on the right and left side of the table an equal number of times. Chimpanzees who passed the seen tracking test within four sessions proceeded to the unseen tracking test.

Unseen tracking test: The unseen tracking test was the true test of chimpanzees' AR-discrimination abilities, because it was impossible for them to succeed in this test by visually tracking the larger grape. As in the previous seen tracking test, chimpanzees witnessed the baiting and the stacking of the containers in the middle of the table, but in this case they were not allowed to witness the positioning of the containers. Rather, after E had baited the containers, she used an opaque occluder to block the chimpanzee's view as she placed the containers on the left and right sides of the table (again, with the position of the larger grape randomly determined). E then removed the occluder and pushed the containers forward. In order to choose correctly in this test, chimpanzees needed to recognize that the grape that looked smaller was actually the larger grape. Chimpanzees were given a maximum of four sessions to pass this test, and those who passed proceeded to the avoid-lens control test.

2.1.4.4. Avoid-lens test: The purpose of this final control test was to determine whether chimpanzees who had passed the previous tests had done so by learning to identify and avoid the magnifying lens container, despite our best efforts to make the containers identical. This control test would also reveal if they had merely learned to look for a specific size of grape and choose that container. The procedure was similar to the previous unseen tracking test, with three changes: (1) instead of using a magnifying lens and a minimizing lens, we used a magnifying lens and a regular, non-distorting piece of glass; (2) we placed the large grape behind the magnifying lens, making it appear even larger than it actually was; and (3) we placed a medium-sized grape behind the regular glass, so that it appeared the same size that the smaller grape had appeared behind the magnifying lens in the experimental trials. The result was that chimpanzees were now faced with a situation in which to choose the truly larger grape, they had to choose (rather than avoid) the magnifying lens. They also had to forego the grape that looked the same size as that which had been correct in the experimental trials and choose the other one instead. Chimpanzees were given a maximum of four sessions to pass this test.

2.2. Results and Discussion

2.2.1. Criteria for passing—Given the small number of subjects and the stepwise nature of the tests, we focused on individual rather than group performance. Chimpanzees were considered to have passed a test if they chose the truly larger grape significantly more often than 50% chance within the maximum number of sessions allowed. We used the binomial test to assess significance, and the passing criteria were as follows: at least 10/12 correct trials in one session ($p = .019$) or at least 8/12 correct trials in each of two consecutive sessions ($p = .038$).

2.2.2. Preference tests—All five chimpanzees passed the preference tests within two sessions (three of them within the first session), clearly demonstrating a natural preference for larger grapes over smaller grapes.

2.2.3. Appearance-reality (AR) test—All five chimpanzees passed the basic test within the maximum allowable eight sessions (see Figure 3). All of the chimpanzees thus demonstrated that they were capable of ignoring the misleading appearance of the grapes and choosing the one that was truly larger.

In addition, every chimpanzee passed the seen and unseen tracking tests and the avoid-lens test within the maximum four sessions. In doing so, they demonstrated that they did not have to rely on visually tracking the larger grape to pass the test, and they had not simply learned to detect and avoid the magnifying-lens container. It is also highly unlikely that the chimpanzees were simply learning the reverse contingency rule ‘choose the small grape to get the large grape’ given their poor performance in previous studies that have tested the use of such rules. Krachun et al. (2009) ran two reverse contingency tests that were very similar in structure to the lens test described here, yet chimpanzees did very poorly in both tests. Boysen et al. (2001) also found that chimpanzees were strongly biased toward selecting an array of food items in which each individual item was large over an equally numbered array of food items in which each item was small, despite the fact that a reversed reward contingency was in effect.

The standard explanation for why chimpanzees perform so poorly with reverse contingency tests is that they find it exceedingly difficult to inhibit their prepotent tendency to choose the larger food reward. But why, then, do chimpanzees perform better in the AR lens test, given that both types of tests involve making choices based on visible features of the stimuli? We suggest that chimpanzees are better able to override prepotent response tendencies in the AR test precisely *because* they recognize that the appearance of the grapes is false. That is, in a standard reverse contingency test, chimpanzees are looking at what appears to be, and what they believe to be, an actual large food item. Thus, both the appearance of a large food item, and the accompanying belief that the food item is actually large, conspire to strongly trigger the chimpanzees’ prepotent tendency to reach for it. In the AR test, however, the belief component of this trigger is missing (at least for chimpanzees who can make an AR discrimination), and the prepotent tendency is consequently reduced. That is, in the AR test, chimpanzees do not believe that the grape they are looking at is an actual large grape, but just that it *appears* large. Without the belief that the grape is actually large, the trigger for the chimpanzees’ prepotent tendency to reach for the grape is lessened. Chimpanzees’ discrepant performance in the two types of tests thus supports the conclusion that they are making an AR discrimination in the lens test.

One might also argue that a simple associative account can be given, whereby chimpanzees came to associate the large grape with negative outcomes in the earliest trials and thereafter avoided choosing the large grape. We believe this is an unlikely explanation for the chimpanzees’ performance for three reasons: (1) they continued to choose the large grape during the warmup preference trials at the start of each test session; (2) they had no hesitancy in choosing the grape that appeared larger behind the lens in the avoid-lens test, demonstrating that they had not learned a specific association between large grapes behind lenses and negative outcomes; and (3) if apes could very quickly form such associations between large grapes and negative outcomes in these kinds of situations, this should allow them to quickly master reverse contingency tests involving different sizes of food, which

they cannot do. Having ruled out these alternative explanations, we believe that the best explanation for the chimpanzees' success was that they came to recognize that the lenses changed the appearance of the grapes but not their true properties.

The high success rate for our current lens test was impressive, given that just five out of 14 chimpanzees succeeded in a similar test by Krachun et al. (2009). The reason for this difference was likely due to our giving chimpanzees more opportunities to learn about the properties of the lenses than they had received in Krachun et al. (2009). In that study, chimpanzees had to pass each test within just two sessions (and, further, the tracking test only included 6 seen and 6 unseen tracking trials per session for a total of 12 trials each). Thus, in the current study, chimpanzees had up to four times as many trials to learn about the stimuli and to pass the tests as in Krachun et al. (2009). It is noteworthy that most of our chimpanzees did not require anywhere near the maximum allowable number of trials to succeed (see Table 2). They passed the basic test within an average of 4 sessions (range = 3–7), the seen tracking test within an average of 1.6 sessions (range = 1–3), the unseen tracking test within an average of 2.6 sessions (range = 2–4), and the avoid-stimulus control test within an average of 1.4 sessions (range = 1–3). The chimpanzees' performance on the very first trial of the crucial control tests further supports our claim that the apes were making an AR discrimination, as opposed to tracking the larger grape or learning a reverse contingency rule. As shown in Table 3, three of the five apes responded correctly on the first trial of the unseen tracking test, and four of the five responded correctly on the first trial of the avoid-stimulus (lens) test.

3. Experiment 2: Mirror Test

In Experiment 1, we were successful in replicating Krachun et al.'s (2009) results with a new population of chimpanzees. Furthermore, we showed that the success rate in the lens test could be vastly improved by giving chimpanzees sufficient time and experience to learn about the illusory properties of the lenses. In Experiment 2, our aim was to investigate whether chimpanzees were capable of passing an appearance-reality test involving a number illusion rather than a size illusion. In this experiment, the chimpanzees needed to recognize that what appeared to be the larger number of food items was, in reality, the smaller number.

3.1. Method

3.1.1. Subjects—Six adult apes from research facilities in Atlanta participated in this experiment, two from the Yerkes National Primate Research Center (Yerkes) and four from Georgia State University's Language Research Center (LRC) (see Table 1).¹ The Yerkes chimpanzees were tested in their outdoor enclosures, and the LRC chimpanzees were tested in their indoor living and sleeping quarters. Subjects were not food-deprived, had water available *ad libitum*, and participated voluntarily in all sessions by approaching the experimenter when called. Three of the four LRC chimpanzees (Lana, Panzee, and Sherman) had, to varying degrees, been previously trained on an artificial communication

¹Three additional chimpanzees from Yerkes (Brandy, Duff and Katrina) failed to reach criterion with the maximum number of preference test sessions and therefore could not participate in the mirror test. A fourth individual at Yerkes (Scott) passed the preference tests but had to be dropped from testing after his fifth basic test session because he repeatedly failed the warmup preference trials.

system involving lexigrams (Savage-Rumbaugh, Shanker, & Taylor, 1998) and had also participated in number discrimination tests (Beran, 2001; Beran & Beran, 2004; Rumbaugh, Savage-Rumbaugh & Hegel, 1987). Prior to the mirror test, none of the LRC chimpanzees had any previous testing experience with AR discrimination tasks. The two Yerkes chimpanzees (Lucas and Faye) had previously participated in at least one of our other AR tests prior to the mirror test (see Table 1).

3.1.2. Design—Experiment 2 included five tests presented in a fixed order: (1) two preference tests, one for each relevant number combination (see below) – each preference test included a maximum of four 12-trial sessions; (2) the basic test with a maximum of eight 12-trial sessions; (3) the seen tracking test with a maximum of four 12-trial sessions; (4) the unseen tracking test with a maximum of four 12-trial sessions; and (5) the avoid-mirror test with a maximum of four 12-trial sessions. For each of the tests (excluding the preference tests), each session began with 2–4 warmup preference trials and 2 demo trials, administered just before the 12 test trials. In every test, the larger number of grapes was on the left and right side an equal number of times, its position was randomized across trials, and it was never in the same position on more than two consecutive trials. Testing proceeded in a stepwise fashion, such that chimpanzees needed to pass one test in order to advance to the next one in the sequence. Passing a test was defined as choosing the larger number of grapes significantly more often than chance within the maximum number of sessions allowed for that test (see results section for details).

3.1.3. Experimental stimuli—The illusory stimuli for this study consisted of two rectangular plastic boxes (21.5 cm X 16.5 cm X 11 cm) into which a mirror (16 cm X 11 cm) could be vertically mounted, near the middle of the box. When the mirror was placed into the box, the grapes in front of it appeared to double in number (see Figure 1b). Each box also included a vertical ‘collar’ (27.5 cm X 23 cm) that slid down over the middle of the box. The collars obscured chimpanzees’ view of the back section of the boxes and also the top edge of the mirror. With the collars in place, it was virtually impossible to tell which box had the mirror inside and which was the regular box (confirmed by several researchers). During testing, we angled the boxes downward slightly by attaching a 1-inch block to the bottom of each box near the back end. This made it possible for chimpanzees to see inside the front of the boxes without being able to see their own reflection in the mirror.

3.1.4. Procedure—All chimpanzees were tested individually. During testing, they sat facing the experimenter (E) across a table (height = 50 cm) positioned in front of their enclosure. The stimuli were displayed on the table (approximately 30 cm apart), and chimpanzees could make a choice by pointing to one of the containers.

3.1.4.1. Preference tests: We first needed to confirm that the chimpanzees preferred larger numbers of grapes to smaller numbers. This was done by presenting chimpanzees with a simple choice between relatively larger versus smaller groups of grapes displayed on separate trays on the table top. Chimpanzees received two preference tests, with the number combinations we presented to them reflecting the choices they would encounter in the experimental trials. They would be choosing between 2 vs 3 grapes in the experimental

trials, and this would look like a choice between 3 vs 4 grapes due to the mirror. Thus, in their first preference test chimpanzees were given a choice between 2 vs 3 grapes, and in their second preference test they were given a choice between 3 vs 4 grapes.² Chimpanzees who showed a statistically significant preference for the larger number of grapes within a maximum of four 12-trial sessions for each of the two preference tests participated in further testing.

Once testing proper began, two preference trials were repeated at the start of each session, as a warmup and also to confirm that the chimpanzees were motivated enough to participate at that time. The number combination presented in these warmup trials was equivalent to how the grapes would *appear* in the experimental trials (e.g., 3 vs 4). When chimpanzees showed a preference for the larger number of grapes in both of these warmup trials, we proceeded with the test session. When they did not do so, we gave them two additional trials and proceeded with testing only if they chose the larger number in at least three of the four trials. Otherwise, the session was postponed until later (this occurred just five times across all five subjects).

3.1.4.2. Basic test: The basic mirror test investigated whether chimpanzees were capable of optimizing their reward by ignoring the apparent number of grapes in each group and choosing the group with the truly larger number. Each basic session included the warmup preference trials (see above), two demo trials (described below), and 12 basic trials.

Demo trials: The demo trials were used to teach chimpanzees about the properties of the mirror. E baited the boxes with the two groups of grapes and then placed the mirror into the box containing fewer grapes. E slid the mirror up and down several times so that the chimpanzee could see how it made the grapes in the box appear to double in number. She then placed the collars on the boxes and pushed the boxes forward so that the chimpanzee could choose between them. Once the chimpanzee had made a choice, E demonstrated the illusory properties of the boxes. She first removed the collars from the boxes so that the mirror could now clearly be seen. As the chimpanzee watched, E slowly removed the mirror from the mirror box and replaced it again, so that the chimpanzee could see how the mirror made the box appear to have more grapes in it than it actually had. E did this at least twice, then removed the mirror and placed it aside. During this demonstration, E also performed other actions to help the chimpanzee learn about the mirrors, for example sweeping her hand through the non-mirror box to show that it did not contain a mirror, and tilting the mirror box forward so that the chimpanzee could look inside and see that it was empty at the back. When E had finished the demonstration, she handed the chimpanzee the grapes from the chosen box and discarded the other grapes. In all test sessions, two such demo trials were repeated after the warmup preference trials, to remind chimpanzees about the illusory properties of the mirror.

²For two chimpanzees, Faye and Scott, we used different number combinations in the basic test (namely, 3 vs 5, which looked like 5 vs 6 with the mirror in place). These two chimpanzees were thus given preference tests for these two number combinations, and their warmup preference trials consisted of the 5 vs 6 combination. We tried this procedural modification with these chimpanzees because we thought that increasing the cost of choosing wrongly might improve performance in the mirror test. Nevertheless, Faye failed the basic test and Scott was unable to complete it due to repeatedly failing the warmup trials.

Basic trials: The basic trials were similar to the previous demo trials, except that E did not perform an extensive demonstration of the properties of the mirror. In the basic trials, E placed the two boxes on the table side by side. (The mirror box did not, at this point, have the mirror inside.) She then placed the two groups of grapes, one smaller and one larger, next to each other on the table in front of the boxes. E then baited the boxes, putting one group of grapes in each box. As the chimpanzee watched, E inserted the mirror into the box containing the smaller number of grapes, so that it now appeared to contain the larger number. She then placed the collars onto the boxes and pushed the containers apart and forward on the table. Once the chimpanzee had chosen a box, E removed the collars, removed the mirror, handed the chimpanzee the grapes from the chosen box, and discarded the other grapes. Chimpanzees were given a maximum of eight sessions to pass the basic test.

3.1.4.3. Tracking control test: Chimpanzees who passed the basic test were next tested to see if they could succeed when they were no longer able to visually track the location of the truly larger number of grapes (see Figure 2). As in Experiment 1, we tested this by first giving chimpanzees a transitional seen tracking test to get them used to the new stacking procedure, followed by the unseen tracking test.

Seen tracking test: In the seen tracking test, E placed the two groups of grapes on the table. She placed one box in the middle of the table and baited it with the larger number of grapes, then stacked the second box on top of the first and baited it with the smaller number of grapes. E then placed the mirror into the top box, just behind the grapes, and placed the collar onto this box. Finally, E unstacked the boxes, placed the collar onto the other box, and pushed the boxes apart and forward. The box with the truly larger number of grapes inside was randomly determined but was on the left and right side of the table an equal number of times. When the chimpanzee had chosen a box, E removed the collars from both boxes, removed the mirror, gave the chimpanzee the grapes from the chosen box, and discarded the other grapes. Chimpanzees who passed the seen tracking test within four sessions advanced to the unseen tracking test.

Unseen tracking test: The procedure for this test was the same as for the seen tracking test except that after baiting both boxes and inserting the mirror E raised an occluder to block the chimpanzee's view. Thus, the chimpanzee saw the initial baiting of the boxes and the placement of the mirror, but did not see the unstacking of the boxes, the placement of collars, and the positioning of boxes on the table. Once the boxes were positioned (with position determined randomly), E removed the occluder and pushed the boxes forward. As before, once the chimpanzee had chosen a box, E removed the collars from both boxes, removed the mirror, gave the chimpanzee the grapes from the chosen box, and discarded the other grapes. Chimpanzees had up to four sessions to pass the unseen tracking test.

3.1.4.4. Avoid-mirror test: Chimpanzees who passed the unseen tracking test were given a final control test to determine whether they had passed the mirror test by learning to somehow detect and avoid the mirror. This control test would also reveal if chimpanzees had merely learned to look for a specific number of grapes (e.g., four) and avoid that group due

to its association with unfavorable outcomes. The procedure was the same as in the previous unseen tracking test, except that E now used a 1 vs 2 grape comparison, putting the single grape into the non-mirror box and two grapes into the mirror box, so that it looked like a choice between 1 vs 4 grapes. Thus, chimpanzees now had to choose (rather than avoid) the mirror box in order to get the truly larger group of two grapes. Chimpanzees also had to choose what looked like four grapes, which had been the incorrect choice throughout all of the experimental trials up to that point. Chimpanzees who passed this control test demonstrated that they were not relying on the simple strategy of avoiding the mirror or the four-grape configuration to pass the AR test.

3.2. Results and Discussion

3.2.1. Criteria for passing—As in Experiment 1, the binomial test was used to assess individual performance and the same criteria were used for passing: at least 10/12 correct trials in one session ($p = .019$) or at least 8/12 correct trials in each of two consecutive sessions ($p = .038$).

3.2.2. Preference tests—As noted in the procedure above, two preference tests were given to each chimpanzee, the first test reflecting the true numbers they would be choosing between in the basic test trials, and the second test reflecting the apparent numbers they would be choosing between. Three chimpanzees who received preference testing did not reach criterion within the four-session maximum and therefore could not participate in the mirror test. However, chimpanzees who passed the preference tests did so quickly (in an average of 1.33 sessions for the first preference test and 1.5 sessions for the second preference test). These chimpanzees clearly demonstrated that they had a natural preference for the larger number of grapes.

3.2.3. Appearance-reality (AR) test—Three of the six chimpanzees tested passed the basic test with mirrors (Figure 3), demonstrating an ability to ignore the illusory appearance of the grapes and to choose the truly larger number of grapes. One individual did so within two sessions, one within five sessions, and one within six sessions (Table 4). All three of those chimpanzees also went on to pass the seen tracking test within one or two sessions.

One chimpanzee, Panzee, also passed the unseen tracking test, and she did so in just one session with a score of 10/12. Panzee then went on to earn a perfect score of 12/12 in the avoid-mirror control test. Further, in each of these control tests she responded correctly from the very first trial (Table 3). Panzee's performance in the avoid-mirror control test demonstrated several things: she had not simply learned to avoid the mirror; she was not just avoiding the specific configuration of four grapes, and; she had not just learned to apply a general reverse contingency rule. Further supporting the last point are results from past studies, in which chimpanzees have had a great degree of difficulty learning to choose a smaller number of food items to obtain the larger number of items (Boysen & Berntson, 1995; Boysen et al., 1996, 1999, 2001; Uher & Call, 2008; Vlamings et al., 2006).

Why did most of the chimpanzees tested find the mirror test more challenging than the lens test? One possible explanation is that it took much longer to administer the mirror test than the lens test, as several grapes (rather than one) had to be placed into each container, the

mirror needed to be inserted, and the collars needed to be fitted onto the boxes. These extra steps added up over trials to result in a longer test that may have taxed the chimpanzees' attention. That the chimpanzees were choosing between different numbers of items, rather than between individual grapes of different sizes, could have also been a factor. In contrast to the preference test for the lens study, in which all the chimpanzees easily and quickly passed, three of the nine individuals who were administered the preference tests for the mirror test failed to reach criterion. This suggests that for some chimpanzees, discriminating between a smaller and larger food item may be easier than discriminating between different quantities of food items. Supporting this is the fact that all three individuals who passed the basic test had prior experience with making number discriminations whereas those who failed it had no prior experience. Memory may have also come into play, with the chimpanzees having a harder time remembering which number was the true number, rather than simply remembering that the grape that looked larger was actually smaller, as they could do in the lens test. The highly compelling nature of the mirror illusion may have contributed to this. (Even the experimenters, during moments of distraction while setting up the stimuli, occasionally found themselves reaching for the illusory grapes behind the mirror!) Making the mirror box more easily identifiable (e.g., by dispensing with the collars and leaving the top edge of the mirror visible) may have helped the chimpanzees avoid such mistakes. However, doing so would have introduced the problem that the chimpanzees could learn to avoid the mirror box rather than make a true AR discrimination.

In summary, chimpanzees were less successful in the mirror test than in the lens test, perhaps because some of the task factors noted above made it a more challenging test for the chimpanzees. Nevertheless, one chimpanzee, Panzee, did succeed without effort, passing both the basic and seen tracking tests within just two sessions each, and the unseen tracking and avoid-mirror control tests within just one session each. Panzee's success demonstrated that it is within the capacity of this species to master this particular type of AR test. Perhaps giving chimpanzees more practice making number discriminations, and also making the test shorter (by reducing the number of trials per session and/or streamlining the procedure), could increase success rates in future studies.

4. EXPERIMENT 3: COLOR TEST

In this final experiment, we tested chimpanzees' AR discrimination abilities with illusory color stimuli. As noted in the introduction (section 1), the use of color stimuli has long been a staple procedure employed in developmental science to test children's AR discriminatory abilities (Flavell et al., 1983; Gopnik & Astington, 1988; Moll & Meltzoff, 2011). As chimpanzees also have trichromatic color vision (Grether, 1940), procedures using color stimuli can be adapted for testing AR discrimination in these animals. Thus, to further investigate the range of chimpanzees' AR discriminatory abilities in comparison to children, Experiment 3 tested chimpanzees' ability to discriminate the apparent color of a stimulus (a box) from its real color.

4.1. Method

4.1.1. Subjects—Seven adult chimpanzees housed at research facilities in Atlanta, Georgia participated, four from the Yerkes National Primate Research Center (Yerkes) and

three from the Language Research Center (LRC) (see Table 1). The Yerkes chimpanzees were tested in their outdoor enclosures and the LRC chimpanzees were tested in their indoor sleeping quarters. Subjects were not food deprived, had water available *ad libitum*, and participated voluntarily in the experiment. All but one of the chimpanzees (Cissie) had previously participated in Experiment 1 and/or 2; and three of them (Cissie, Faye, and Lucas) had participated in an earlier pilot study using small color filters. None of the subjects in the pilot study, however, succeeded in passing the tracking test. We suspected that the diminutive size of the filters (4×4 cm) was a major reason for the difficulty, and we thus created larger color filters (15×15 cm) to be used in the current experiment.

4.1.2. Design—Experiment 3 included five tests presented in a fixed order: (1) a two-step pretest training phase (each step consisting of a maximum of eight 16-trial sessions), in which chimpanzees learned that the white box was always empty but the color (green or yellow) box always contained a piece of banana; (2) the basic test with a maximum of eight 16-trial sessions; (3) the seen tracking test with a maximum of four 16-trial sessions; (4) the unseen tracking test with a maximum of four 16-trial sessions; and (5) a two-step avoid-filter test with a maximum of four 16-trial sessions per step. For each of the tests (excluding the training sessions), each session included 2 demo trials, administered immediately before the 16 test trials. In every session, the baited box was the green one in half the trials and the yellow one in half the trials. In addition, the baited box was on the left and right side an equal number of times, its position was randomized across trials, and it was never in the same position for more than two consecutive trials. Testing proceeded in a stepwise fashion, and chimpanzees needed to pass one test in order to advance to the next. Passing a test was defined as choosing the color box significantly more often than chance within the maximum number of sessions allowed for that test (see results section for details).

4.1.3. Experimental stimuli—The stimuli for this study consisted of three cardboard boxes ($9 \times 9 \times 9$ cm) that were identical except for their color: white, green or yellow. Also used were three transparent filters (15×15 cm) mounted into frames constructed of plastic: an untinted ('clear') filter, a green filter, and a yellow filter. The height and width of the filters were larger than the height and width of the front of the boxes, to allow the filters to extend out beyond the front of the boxes once the boxes were placed behind them (see Figure 1c). This construction of the stimuli matches the general design used in children's studies (e.g., Flavell et al., 1983; Moll & Meltzoff, 2011). The colors of the filters were such that when the white box was placed behind the yellow filter it looked the same color as the truly yellow box, and when it was placed behind the green filter it looked the same color as the truly green box.

4.1.4. Procedure—All chimpanzees were tested individually. During testing, they sat facing an experimenter (E) across a table (height = 50 cm) positioned in front of their enclosure. The stimuli were displayed on the table (approximately 30 cm apart), and chimpanzees could make a choice by pointing to one of the containers.

4.1.4.1. Pretest training: In preparation for testing, it was necessary for chimpanzees to learn that the white box was always empty while both color (green and yellow) boxes always contained a piece of banana. The training phase involved two steps, as follows:

Training step 1: In this first step, the experimenter (E) sat opposite the chimpanzee and placed a white box and a color box on the table top. She allowed the chimpanzee to watch as she hid a piece of banana inside the color box, and she also showed the chimpanzee that the white box was empty. E then pushed the boxes forward. When chimpanzees chose the color box, E gave them the banana inside; when they chose the white box, she showed them that the box was empty and then removed the banana from color box and discarded it. Chimpanzees were allowed a maximum of eight sessions, each containing 16 trials (8 trials with a white box vs a yellow box, and 8 trials with a white box vs a green box, randomly intermixed). Chimpanzees passed this stage of training when they were choosing the color box over the white box significantly more often than chance.

Training step 2: This second step differed from the first only in that E did not allow chimpanzees to see which box she put the banana into; they thus had to learn that the banana would always be in the color box and never in the white box. The procedure was as follows: E placed a white box and a color box on the table top. She blocked the chimpanzee's view with an opaque occluder and baited the color box. E then removed the occluder and pushed the boxes forward on the table so that the chimpanzee could make a choice. Chimpanzees needed to choose the color box rather than the white box to get the banana hidden inside. Again, chimpanzees were allowed a maximum of eight sessions to pass, and they had to simultaneously pass for each color combination to move on to testing.

4.1.4.2. Basic color test: The basic color test consisted of 2 demo trials followed by 16 basic trials (8 for each color illusion).

Demo trials: Each basic test session began with two demo trials. The demo trials served as a warmup and gave chimpanzees the opportunity to learn about the illusory properties of the color filters. In one of the demo trials, chimpanzees were given a choice between a green box behind a clear filter so that it still looked green, and a white box behind a yellow filter so that it now looked yellow (called the G-G vs W-Y combo). In the other demo trial, they were given a choice between a yellow box behind a clear filter so that it still looked yellow, and a white box behind a green filter so that it now looked green (called the Y-Y vs W-G combo). The order of these trials was counterbalanced across subjects. The trials proceeded as follows: In the G-G vs W-Y combo trials, E placed a clear filter and a yellow filter on the table. She also placed a green box on the table next to the clear filter and a white box on the table next to the yellow filter. E showed the chimpanzee that she had a piece of banana and then hid the banana inside the green box. After the box was baited, E slowly placed the green box behind the clear filter (it still appeared green), and the white box behind the yellow filter (it now appeared yellow). E moved the white box behind the yellow filter and back out again at least twice to give the chimpanzee ample opportunity to see the change in the appearance of the box. E then pushed the stimuli apart and forward so that the chimpanzee could choose between them. If chimpanzees were able to recognize that the box

behind the yellow filter was apparently yellow but truly white, they should choose the green box behind the clear filter. When they did so, E removed the banana and gave it to the chimpanzee. When they instead chose the white box, E slowly removed that box from behind the yellow filter and showed the chimpanzee that it was empty. She then removed the green box from behind the clear filter, removed the banana from that box, and discarded it. Trials for the Y-Y vs W-G combination followed the same procedure, except that the yellow box was placed behind the clear filter and the white box was placed behind the green filter.

Basic trials: In the basic trials, we tested chimpanzees' ability to make a choice based on the actual, rather than apparent, colors of the boxes. Again, chimpanzees received two types of trials: G-G vs W-Y trials and Y-Y vs W-G trials (8 of each per session). The procedure for these trials was similar to the previous demo trials but varied in some important ways. In the G-G vs W-Y trials, E placed a clear filter and a yellow filter on the table. She then placed a green box on the table next to the clear filter and white box on the table next to the yellow filter, as in the previous demo trials. However, unlike in the demo trials, E blocked the chimpanzee's view with an opaque occluder before baiting the green box. Thus, chimpanzees were aware of the fact that E was baiting one of the boxes, but they had to rely on what they learned during training to infer that she was baiting the green (rather than white) box. Once the box was baited, E removed the occluder and allowed the chimpanzee to watch as she placed the green box behind the clear filter (it still appeared green), and the white box behind the yellow filter (it now appeared yellow). E then pushed the stimuli apart and forward so that the chimpanzee could choose between them. Chimpanzees who were able to recognize that the box behind the yellow filter was apparently yellow but truly white would be expected to choose the truly green box behind the clear filter to get the hidden reward. The Y-Y vs W-G trials followed the same procedure, except that the yellow box was placed behind the clear filter and the white box was placed behind the green filter. In those trials, the correct choice was to choose the truly yellow box behind the clear filter to get the reward. Chimpanzees were given a maximum of eight sessions of the basic test. To pass, they had to achieve significantly better than chance performance simultaneously for both color combinations by choosing the truly yellow or green box over the truly white box.

4.1.4.3. Tracking control test: Chimpanzees who passed the basic test were next tested to see if they could succeed when they were no longer able to visually track the location of the actual color box. As in the previous two experiments, we began with a transitional seen tracking test to familiarize the chimpanzees with the stacking procedure, and we then ran the diagnostic unseen tracking test.

Seen tracking test: The seen tracking test was similar to the basic test trials except that after E placed the boxes behind the filters, she briefly stacked the stimuli in the center of the table before positioning them on the right and left sides of the table (position determined randomly) and then pushing them forward so that the chimpanzee could choose. Chimpanzees were given a maximum of four 16-trial sessions, and they had to achieve significantly better than chance performance simultaneously for both color combinations before moving on to the unseen tracking test.

Unseen tracking test: The unseen tracking test differed from the seen tracking test in that after E stacked the stimuli in the center of the table, she occluded the chimpanzee's view as she positioned the stimuli on the left and right sides of the table (position determined randomly). She then removed the occluder and pushed the stimuli forward so that the chimpanzee could choose. Thus, in the unseen tracking test, chimpanzees had to take note of which of the two boxes was the color box and which was the white box while E was placing them behind the filters, and then look for that box later when it was time to make a choice—they could not simply visually track the location of the boxes. Chimpanzees were given a maximum of four 16-trial sessions of the unseen tracking test. To pass, they had to achieve significantly better than chance performance simultaneously for both color combinations by choosing the actual color box over the white box.

4.1.4.4. Avoid-filter test: As in the previous two experiments, we also wanted to rule out the possibility that chimpanzees who succeeded in our test were doing so by learning to avoid the illusion-producing stimulus. In this experiment, the concern was that chimpanzees could simply look for the color filter—whether yellow or green—and avoid it. Alternatively, they could look for the clear filter and choose that one. We tested this possibility by presenting chimpanzees with a situation in which both filters on the table were color filters. The color combinations used were still G-G vs W-Y and Y-Y vs W-G, but instead of placing the color box behind a clear filter, we placed it behind a filter that was the same color as the box itself: green box behind green filter, and yellow box behind yellow filter. (We adjusted the color saturations so that the boxes, when placed behind the color filters, appeared the same color as when placed behind the clear filters during the previous experimental trials.) Because we were introducing the chimpanzees to a new situation in this control test—two color filters present instead of just one—we carried out the test in two steps, an *unstacked* version followed by a *stacked* version.

Unstacked version: Behind an occluder and out of the chimpanzee's view, E placed a green and a yellow filter as well as a white box and a color box on the table, and she baited the color box. E then removed the occluder and, as the chimpanzee watched, placed the color box behind the filter of the same color. She also put the white box behind the other color filter. E then pushed the boxes apart and forward on the table. When chimpanzees chose the actual color box, E removed the banana and gave it them. When they chose the truly white box E removed the box from behind the filter and revealed that it was empty. She then removed the actual color box from behind its color filter and discarded the banana. Chimpanzees were given a maximum of four sessions to pass the unstacked version of this test. Those who passed (for each color combination separately) moved on to the more challenging stacked test in which they could not visually track the boxes.

Stacked version: In this second step of the avoid-filter test, we tested whether chimpanzees could pass the test when we used a more challenging procedure that involved stacking the stimuli and positioning them on the table out of their view (as in the previous unseen tracking control test described above). This stacked version of the avoid-filter test initially proceeded in the same manner as the unstacked version above. Behind an occluder, E placed the filters (green and yellow) and boxes (one color box and one white box) on the table, and

then baited the color box. She then removed the occluder and placed the boxes behind the filters (color box behind the same-color filter, white box behind the other color filter). At this point, however, rather than pushing the stimuli apart and forward, E stacked one on top of the other and then occluded the chimpanzee's view while she positioned the stimuli on the left and right sides of the table. E then removed the occluder and pushed the stimuli forward. As before, when chimpanzees chose the actual color box, they were given the banana inside; when they chose the truly white box E removed the boxes from behind the filters, revealed that the white box was empty, and discarded the banana from the other box.

4.2. Results and Discussion

4.2.1. Criteria for passing—There were 16 test trials per session in this experiment, 8 trials for each color combination: G-G vs W-Y, and Y-Y vs W-G. At any given stage of testing, chimpanzees had to be at criterion for both color combinations separately within the same session to pass. The criteria for passing, based on the binomial test, were as follows: at least 7/8 correct in one session ($p = .035$), or at least 6/8 correct in each of two consecutive sessions ($p = .021$).

4.2.2. Pretest training—All chimpanzees tested were capable of learning that the white box was always empty and the color boxes always contained a piece of banana. For training step 1, the mean number of sessions to reach criterion was 1.43 (range = 1–3). For training step 2, the mean number of sessions to criterion was 1.85 (range = 1–2).

4.2.3. Appearance-reality (AR) test—Five of the seven chimpanzees tested passed the basic color test (Figure 3), taking an average of 3.2 sessions (range = 1–5) to do so (Table 5). Those five chimpanzees also passed the seen tracking test (mean number of sessions = 1.6 sessions, range = 1–2), which differed from the basic test only in that E briefly stacked the stimuli in the center of the table just before positioning them on the left and right sides of the table. Thus, most chimpanzees demonstrated that they were capable of ignoring the illusory appearance of the white box behind the color filter and choosing the truly green/yellow box instead.

Two chimpanzees (Lamar and Lucas) passed the unseen tracking test in 2 and 4 sessions, respectively, and they both also responded correctly in the first trial of that test (Table 3). This allows us to say with certainty that those two chimpanzees did not rely on visual tracking of the actual color box to pass the test. Lamar was unfortunately unavailable to participate in the avoid-filter test, but Lucas did pass, taking 2 sessions for each of the unstacked and stacked versions. Although Lucas did not respond correctly in the very first trial of the avoid-filter test, he did respond correctly in the three subsequent trials, and his overall score for the first session of that test was 13/16 trials correct (six of which occurred in the first half of the session). Thus, Lucas' performance in the avoid-filter test demonstrated that he was not relying on the simple strategy of avoiding the color filter (or looking for the clear filter) to pass the test.

5. GENERAL DISCUSSION

Krachun et al. (2009) previously demonstrated that chimpanzees were capable of passing an appearance-reality test in which distorting lenses were used to change the apparent size of food items. In Experiment 1 of the current study, we replicated this result with a different population of chimpanzees, and we showed that giving chimpanzees more opportunities to learn about the illusory properties of the lenses could greatly increase success rates. More notably, in Experiments 2 and 3, we found that some chimpanzees were also capable of making AR discriminations when the properties manipulated were the apparent number or color of the objects. These findings widen the scope of chimpanzees' AR discrimination capacities, suggesting that they may possess a general understanding that the perceptual appearance of objects can, in a variety of ways, misrepresent their true properties. We do not want to overstate the case, given that only a few of the chimpanzees who passed the basic mirror and color tests also went on to pass the subsequent (and crucial) unseen tracking control test. However, as noted earlier in our discussion of the mirror test results, some features of the tests that were unrelated to the AR discrimination question (e.g., the necessity of attending to a relatively lengthy sequence of events) could have added to their difficulty. Individual differences in concentration, motivation, and short-term memory may have thus played a significant role in deciding which individuals would be successful in the more challenging tracking control test. While the consequence of this is that our results might underestimate chimpanzees' AR discrimination abilities, we believed the tracking control test was essential to rule out the possibility that our subjects were merely visually tracking the desired stimulus. The challenge for future research is to design AR tests that achieve adequate experimental control without taxing animals' cognitive resources in ways that interfere with their ability to demonstrate the full extent of their AR discrimination abilities. In addition, future research into the range of apes' capacities would be greatly benefited by examining their AR discrimination abilities for identity illusions as well as property illusions.

The results of our three studies suggest that, for a range of properties, chimpanzees are capable of distinguishing apparent from real properties of stimuli, and can use the apparent properties (e.g., the apparent small size of a grape) to locate the stimulus with the truly desirable property (e.g., the larger grape). However, three alternatives to this appearance-reality discrimination explanation need to be considered. First, it could be suggested that the chimpanzees solved our tasks not by discriminating the apparent properties from the real properties of the stimuli, but by quickly learning a reverse contingency rule, such as 'to get the larger grape (or larger number of grapes) choose the container with the smaller grape (or fewer number of grapes).' We did not run reverse contingency control tests in this study because, as noted earlier, previous such tests (Boysen & Berntson, 1995; Boysen et al., 1996, 1999, 2001; Krachun et al., 2009; Uher & Call, 2008; Vlamings et al., 2006) have firmly demonstrated that chimpanzees have great difficulty learning such rules for both the size and quantity of desirable food items. Thus, the quick learning of a reverse contingency rule is a very unlikely explanation for successful chimpanzees' performance in Experiments 1 and 2. This is most clearly evident in the case of Experiment 2, in which chimpanzees were required to discriminate the apparent from real quantity of grapes. As noted above, a

number of previous studies show that chimpanzees find it exceedingly difficult to learn reverse contingency rules for quantity; however, the chimpanzee that passed our Mirror Test did so quite easily, requiring no more than two sessions (24 trials) to pass each stage of testing. In addition, the results of the avoid-stimulus control tests of those experiments (i.e., the avoid-lens and avoid-mirror tests, respectively) go against the reverse contingency explanation, since chimpanzees consistently chose the apparently larger grape or larger number of grapes in those tests (note that they also did so in the warmup preference trials at the start of each test session). Finally, the reverse contingency explanation does not extend to Experiment 3, which did not include any apparent reverse contingency.

The second possible alternative explanation refers specifically to chimpanzees' performance in our unseen tracking control tests. The 'image tracking hypothesis' (proposed by Carruthers, in conversation; see also Lurz, 2011, p. 208) suggests that through their experiences with the illusory properties of the lens containers and the mirror box, chimpanzees come to interpret these objects as forming images. From their experience with the lenses, for example, they come to believe that when the large grape is placed behind the minimizing lens, the lens projects a small-grape image on its surface. Thus, it is this image that chimpanzees take themselves to be looking at when they look at the baited container, much as we take ourselves to be looking at images on TV and movie screens. The chimpanzees then use this small-grape image to track the location of the large actual grape, since the two are perfectly correlated through the experiment. A similar explanation can be given for chimpanzees' performance in the mirror test.

However, although chimpanzees are known to be able to use images to track the location of hidden objects (Boysen & Kuhlmeier, 2002; Menzel et al., 1985), we do not believe the image-tracking explanation provides an adequate alternative to the AR discrimination explanation, for the following reasons. In the unseen tracking test, the lens containers have been unstacked and rearranged on the table behind an occluder, out of chimpanzees' view. Thus, when the occluder is removed, chimpanzees need to locate the container that has the small-grape image on it in order to locate the container with the actual large grape inside. But how would they do this? It might be suggested that they could just look at the two containers and then pick the one that looks like it has a small-grape image on it. But neither container looks like it has an image on it; rather, they both look like what they are, containers with different sized grapes inside (which is precisely why chimpanzees typically choose the magnified grape before they have had a chance to learn about the properties of the lenses). It is also important to note that the image-tracking explanation is even less plausible in the case of the chimpanzees' performance in the color test, as there is no reason to think that chimpanzees interpret the color filters as forming images. Thus, we do not think the image-tracking explanation provides a credible alternative to the AR discrimination explanation.

Karg et al. (2014) mention a third alternative explanation for how apes might solve appearance-reality tasks, suggesting that they could employ conservation rules. Such rules could include that properties such as the color, size, and number of objects are conserved (do not change) when they are placed behind a filter or lens, or in front of a mirror. However, although an understanding of such rules is undoubtedly required to pass all of our AR tasks,

it is not sufficient. For in Experiments 1 and 2, the chimpanzees still need to identify the grape or group of grapes whose large size/number has been conserved after it has been placed inside a container, stacked, and positioned on the table behind an occluder. When the occluder is removed, chimpanzees are staring at two containers, one which looks like it has a large grape inside and the other which looks like it has a small grape inside; or one which looks like it has three grapes inside and the other which looks like it has four grapes inside. Knowing the conservation rules above will not help chimpanzees know which grape or group of grapes is the one whose large size or number has been conserved. To know this, chimpanzees must apply some type of appearance-versus-reality rule, such as ‘the grape that appears small is the one whose large size has been conserved,’ or ‘the group of grapes that appears fewer in number is the one whose larger number has been conserved.’ Thus we do not see how conservation abilities, unaided by any understanding of the appearance-reality distinction, would be sufficient to pass the AR tasks in Experiments 1 and 2. We concede that this may not be the case for Experiment 3, in which successful chimpanzees could have arguably passed by using the conservation rule that the white box does not change color when placed behind a color filter. This is because when the stimuli are revealed (after having been stacked and rearranged behind an occluder), chimpanzees only need to remember which of the two color filters (yellow or green) the white box had been placed behind earlier, locate that filter, and avoid selecting it. Although this is a possible explanation for successful performance in Experiment 3, it is also possible that the chimpanzees passed the tests by using appearance-versus-reality rules, such as ‘the box that appears green is really white,’ and ‘the box that appears yellow is really yellow.’ Given that chimpanzees appear to use appearance-versus-reality rules to pass the tests in Experiments 1 and 2, the hypothesis that they also use such rules to pass similar tests in Experiment 3 provides a more unifying explanation of the data from the three experiments than the hypothesis that chimpanzees use quite different strategies to solve similar sets of problems. Furthermore, to pass the tests in Experiment 3, the chimpanzees need to at the very least ignore the deceptive color appearance of the white box and choose the actual color box, which is arguably a type of AR discrimination.

Earlier, we considered the adaptive advantages that AR discrimination abilities might confer on individuals. These advantages could have provided a basis for the eventual instalment of this capacity in the cognitive tool kit of all typically developing humans. That some chimpanzees demonstrate an ability to pass simple physical AR tests implies deep evolutionary roots for our own capacities, which extend far beyond the physical domain. Most adult humans understand that not just objects but also people, animals, emotions, intentions, beliefs, situations, and events can appear to be something other than what they really are. As noted previously, many developmental researchers suggest that when children begin to pass standard AR discrimination tests at 4–5 years of age, it is because they have experienced an important conceptual shift toward understanding the nature of dual representation more generally. This change, they argue, also bears on children’s ability to pass a wide variety of other tests that draw upon the same underlying conceptual proficiency. We suggest that AR discrimination in the physical domain, rather than being just one more manifestation of this conceptual change, could play a more supportive role. An individual who, through experience, becomes aware that objects are not always as they

perceptually appear to be may be primed to understand and predict the actions of others who are fooled by misleading appearances (Lurz, 2011).

If being aware of the appearance-reality distinction provides a foundation for understanding that others' representations may also be at odds with reality, this suggests a direction for future research. We now know that some chimpanzees are capable of passing AR tests. These individuals, by introspecting on their own experiences with misleading perceptions, may have come to understand that other individuals can also have perceptual experiences that do not accord with reality. In human developmental research, children who pass tests requiring insight into their own mistaken perceptions of deceptive objects also expect naïve others to be deceived by the objects (e.g., see Gopnik & Astington, 1988; Moll & Meltzoff, 2011). Thus, an important next step in comparative research will be to devise tests to determine whether this is also the case for chimpanzees.

Acknowledgments

We would like to thank John Kelley, Joe McIntyre, Jennifer Schaeffer, Dan Wilson, and McLennon (Mac) Wilson for their invaluable help with these experiments. We would also like to thank the reviewer and editor for their extremely insightful comments that strengthened the paper. Financial support for this research was provided by a Natural Sciences and Engineering Research Council (NSERC) Discovery Grant (435593-2013) to C. Krachun and a National Institutes of Health (NIH) grant (HD-50563) to W. Hopkins.

References

- Andrews G, Halford G, Bunch K, Bowden D, Jones T. Theory of mind and relational complexity. *Child Development*. 2003; 74:1476–1499.10.1111/1467-8624.00618 [PubMed: 14552409]
- Beran M. Summation and numerosness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*. 2001; 115:181–191.10.1037/0735-7036.115.2.181 [PubMed: 11459165]
- Beran M, Beran MM. Chimpanzees remember the results of one-by-one addition of food items to sets over extended time periods. *Psychological Science*. 2004; 15:94–99.10.1111/j.0963-7214.2004.01502004.x [PubMed: 14738515]
- Boysen ST, Berntson GG. Responses to quantity: Perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*. 1995; 21:76–86.10.1037/0097-7403.21.1.82
- Boysen ST, Berntson GG, Hannan MB, Cacioppo JT. Quantity-based interference and symbolic representations in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*. 1996; 22:76–86.10.1037/0097-7403.22.1.76 [PubMed: 8568498]
- Boysen ST, Berntson GG, Mukobi KL. Size matters: Impact of item size and quantity on array choice by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*. 2001; 115:106–110.10.1037/0735-7036.115.1.106 [PubMed: 11334213]
- Boysen, S.; Kuhlmeier, V. Representational capacities for pretense with scale models and photographs in chimpanzees (*Pan troglodytes*). In: Mitchell, R., editor. *Pretending and imagination in animals and children*. Cambridge: Cambridge University Press; 2002. p. 210-228.
- Boysen ST, Mukobi KL, Berntson GG. Overcoming response bias using symbolic representations of number by chimpanzees (*Pan troglodytes*). *Animal Learning and Behavior*. 1999; 27:229–235.10.3758/BF03199679
- Braine MDS, Shanks BL. The development of conservation of size. *Journal of Verbal Learning and Verbal Behavior*. 1965; 4:227–242.10.1016/S0022-5371(65)80025-1
- Byrne RW, Whiten A. Tactical deception in primates: The 1990 database. *Primate Report*. 1990; 27:1–101.
- Davidson, D. *The problems of rationality*. Oxford: Clarendon Press; 2004.

- Deák GO. Do children really confuse appearance and reality? Trends in Cognitive Sciences. 2006; 10:546–550.10.1016/j.tics.2006.09.012 [PubMed: 17079185]
- Flavell JH. The development of children's knowledge about the appearance-reality distinction. American Psychologist. 1986; 41:418–425.10.1037/0003-066X.41.4.418 [PubMed: 3717749]
- Flavell JH. The development of children's knowledge of false belief and the appearance-reality distinction. International Journal of Psychology. 1993; 28:595–604.10.1080/00207599308246944
- Flavell JH, Flavell ER, Green FL. Development of the appearance-reality distinction. Cognitive Psychology. 1983; 15:95–120.10.1016/0010-0285(83)90005-1 [PubMed: 6831859]
- Flavell JH, Flavell ER, Green FL. Young children's knowledge about the apparent-real and pretend-real distinctions. Developmental Psychology. 1987; 23:816–822.10.1037/0012-1649.23.6.816
- Flavell JH, Green FL, Flavell ER. Development of knowledge about the appearance-reality distinction. Monographs of the Society for Research in Child Development. 1986; 51:i–69.10.2307/1165866 [PubMed: 3807927]
- Friend M, Davis TL. Appearance-reality distinction: Children's understanding of the physical and affective domains. Developmental Psychology. 1993; 29:907–914.10.1037/0012-1649.29.5.907
- Gallup GG Jr. Chimpanzees: Self-recognition. Science. 1970; 167:86–87.10.1126/science.167.3914.86 [PubMed: 4982211]
- Gopnik A, Astington JW. Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. Child Development. 1988; 59:26–37.10.2307/1130386 [PubMed: 3342716]
- Grether WF. A comparison of human and chimpanzee spectral hue discrimination curves. Journal of Experimental Psychology. 1940; 26:394–403.10.1037/h0056423
- Hansen MB, Markman EM. Appearance questions can be misleading: A discourse-based account of the appearance-reality problem. Cognitive Psychology. 2005; 50:233–263.10.1016/j.cogpsych.2004.09.001 [PubMed: 15826611]
- Hare B, Call J, Tomasello M. Chimpanzees deceive a human competitor by hiding. Cognition. 2006; 101:495–514.10.1016/j.cognition.2005.01.011 [PubMed: 16412413]
- Hirata S, Matsuzawa T. Tactics to obtain a hidden food item in chimpanzee pairs (*Pan troglodytes*). Animal Cognition. 2001; 4:285–295.10.1007/s100710100096 [PubMed: 24777519]
- Hopkins WD, Adams M, Weiss A. Genetic and environmental contributions to the expression of handedness in chimpanzees (*Pan troglodytes*). Genes, Brain and Behavior. 2013; 12:446–452.10.1111/gbb.12044
- Hopkins WD, Keebaugh AC, Reamer LA, Schaeffer J, Schapiro SJ, Young LJ. Genetic influences on receptive joint attention in chimpanzees (*Pan troglodytes*). Scientific Reports. 2014; 4:3774.10.1038/srep03774 [PubMed: 24440967]
- Hopkins WD, Reamer L, Mareno MC, Schapiro SJ. Genetic basis for motor skill and hand preference for tool use in chimpanzees (*Pan troglodytes*). Proceedings of the Royal Society of London: Biological Sciences B. 2015; 782:20141223.10.1098/rspb.2014.1223
- Hopkins WD, Russell JL, Schaeffer JS. Chimpanzee intelligence is heritable. Current Biology. 2014; 24:1640–1652.10.1016/j.cub.2014.05.076
- Karg K, Schmelz M, Call J, Tomasello M. All great ape species (*Gorilla gorilla*, *Pan paniscus*, *Pan troglodytes*, *Pongo abelii*) and two-and-a-half-year-old children (*Homo sapiens*) discriminate appearance from reality. Journal of Comparative Psychology. 2014; 128:431–439.10.1037/a0037385 [PubMed: 25150962]
- Krachun C, Call J, Tomasello M. Can chimpanzees (*Pan troglodytes*) discriminate appearance from reality? Cognition. 2009; 112:435–450.10.1016/j.cognition.2009.06.012 [PubMed: 19631933]
- Kuhlmeier VA, Boysen ST, Mukobi KL. Scale-model comprehension by chimpanzees (*Pan troglodytes*). Journal of Comparative Psychology. 1999; 113:396–402.10.1037/0735-7036.113.4.396 [PubMed: 10608563]
- Lurz, R. Mindreading animals. Cambridge, MA: MIT Press; 2011.
- Menzel, EW. A group of young chimpanzees in a one-acre field. In: Schrier, A.; Stollnitz, F., editors. Behavior of nonhuman primates. New York: Academic Press; 1974. p. 83–153.

- Menzel EW, Savage-Rumbaugh ES, Lawson J. Chimpanzee (*Pan troglodytes*) spatial problem solving with the use of mirrors and televised equivalents of mirrors. *Journal of Comparative Psychology*. 1985; 99:211–217.10.1037/0735-7036.99.2.211 [PubMed: 4006435]
- Mitchell, RW.; Thomson, NS., editors. *Deception: Perspectives on human and nonhuman deceit*. Albany, NY: SUNY; 1986.
- Moll H, Meltzoff AN. How does it look? Level 2 perspective-taking at 36 months of age. *Child Development*. 2011; 82:661–673.10.1111/j.1467-8624.2010.01571.x [PubMed: 21410927]
- Moll H, Meltzoff AN, Merzsch K, Tomasello M. Taking versus confronting visual perspectives in preschool children. *Developmental Psychology*. 2012; 49:646–654.10.1037/a0028633 [PubMed: 22612438]
- Muncer S. “Conversations” with a chimpanzee. *Developmental Psychobiology*. 1983; 16:1–11.10.1002/dev.420160102 [PubMed: 6825963]
- Nagel, T. *Mind and cosmos*. Oxford: Oxford University Press; 2012.
- Perner, J. *Understanding the representational mind*. Cambridge, MA: MIT Press; 1991.
- Povinelli DJ, Rulf A, Landau K, Bierschwale D. Self-recognition in chimpanzees (*Pan troglodytes*): Distribution, ontogeny, and patterns of emergence. *Journal of Comparative Psychology*. 1993; 107:347–372.10.1037/0735-7036.107.4.347 [PubMed: 8112048]
- Rice C, Koinis D, Sullivan K, Tager-Flusberg H, Winner E. When 3-year-olds pass the appearance-reality test. *Developmental Psychology*. 1997; 33:54–61.10.1037/0012-1649.33.1.54 [PubMed: 9050390]
- Rumbaugh D, Savage-Rumbaugh S, Hegel M. Summation in the chimpanzee (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*. 1987; 13:107–115.10.1037/0097-7403.13.2.107 [PubMed: 3572305]
- Russell J, Lyn H, Schaeffer JA, Hopkins WD. The role of socio-communicative rearing environments on the development of social and physical cognition in apes. *Developmental Science*. 2011; 14:1459–1470.10.1111/j.1467-7687.2011.01090.x [PubMed: 22010903]
- Sapp F, Lee A, Muir D. Three-year-olds’ difficulty with the appearance-reality distinction: Is it real or is it apparent? *Developmental Psychology*. 2000; 36:547–560.10.1037/0012-1649.36.5.547 [PubMed: 10976596]
- Savage-Rumbaugh, S.; Shanker, SG.; Taylor, TJ. *Apes, language, and the human mind*. Oxford: Oxford University Press; 1998.
- Strawson, PF. *Individuals*. London: Routledge; 1959.
- Suda C, Call J. Piagetian liquid conservation in the great apes (*Pan paniscus*, *Pan troglodytes*, and *Pongo pygmaeus*). *Journal of Comparative Psychology*. 2004; 118:265–279.10.1037/0735-7036.118.3.265 [PubMed: 15482054]
- Taylor M, Hort B. Can children be trained in making the distinction between appearance and reality? *Cognitive Development*. 1990; 5:89–99.10.1016/0885-2014(90)90014-K
- Uher J, Call J. How the great apes (*Pan troglodytes*, *Pongo pygmaeus*, *Pan paniscus*, and *Gorilla gorilla*) perform on the reversed contingency task II: Transfer to new quantities, long-term retention, and the impact of quantity ratios. *Journal of Comparative Psychology*. 2008; 122:204–212.10.1037/0735-7036.122.2.204 [PubMed: 18489236]
- Vlamings P, Uher J, Call J. How the great apes (*Pan troglodytes*, *Pongo pygmaeus*, *Pan paniscus*, and *Gorilla gorilla*) perform on the reversed contingency task: The effects of food quantity and food visibility. *Journal of Experimental Psychology: Animal Behavior Processes*. 2006; 32:60–70.10.1037/0097-7403.32.1.60 [PubMed: 16435965]
- Wellman, HM. *The child’s theory of mind*. MIT Press; Cambridge, MA: 1990.
- Whiten A, Byrne RW. Tactical deception in primates. *Behavioral and Brain Sciences*. 1988; 11:233–273.10.1017/S0140525X00049682
- Woodruff G, Premack D. Intentional communication in the chimpanzee: The development of deception. *Cognition*. 1979; 7:333–362.10.1016/0010-0277(79)90021-0
- Woodruff G, Premack D, Kennel K. Conservation of liquid and solid quantity by the chimpanzee. *Science*. 1978; 202:991–994.10.1126/science.202.4371.991 [PubMed: 17798798]

HIGHLIGHTS

- Novel illusory stimuli were used to test appearance-reality skills in chimpanzees.
- Chimpanzees showed some degree of success in all three tests.
- Control sessions ruled out low-level explanations for their performance.
- Chimpanzees thus demonstrated AR discrimination for a range of object properties.

Figure 1. Illusory stimuli used in the three appearance-reality (AR) tests. Note that for the color test, a green box and a yellow filter were also used. Top row = reality; bottom row = appearance.

Figure 2.

Diagrammed examples of the procedure for the unseen tracking test in each experiment.

After the containers are stacked and then repositioned behind the occluder, they can now be in either their original location or the opposite location (i.e., position is determined randomly). Therefore, chimpanzees cannot just visually track the location of the more desirable option.

Figure 3.

Number of chimpanzees who passed each phase of testing for the three AR tests. Note: in the color test, one chimpanzee who passed the unseen tracking test (Lamar) was unavailable to take the avoid-stimulus test.

Table 1

Chimpanzee participants

Name	Sex	Age	Facility	Test ³
Artemus	M	22	Yerkes ¹	Lens
Cissie	F	37	Yerkes	Color
Faye	F	21	Yerkes	Lens, Mirror, Color
Lamar	M	25	Yerkes	Lens, Color
Lana	F	42	LRC ²	Mirror, Color
Lucas	M	20	Yerkes	Lens, Color, Mirror
Mercury	M	27	LRC	Mirror, Color
Panzee	F	28	LRC	Mirror
Sherman	M	40	LRC	Mirror, Color
Shirley	F	28	Yerkes	Lens

¹Yerkes National Primate Research Center, Emory University

²Language Research Center, Georgia State University

³For each subject, tests are listed in the order in which they were administered.

Lens test – number of trials (out of 12) in which chimpanzees chose the truly larger grape, across all experimental sessions (S)

Table 2

Subject	Basic lens test								Seen tracking			Unseen tracking			Avoid-lens		
	S1	S2	S3	S4	S5	S6	S7	S8	S1	S2	S3	S1	S2	S3	S1	S2	S3
Artemus	5	9	<u>8</u>	.	.	.	.	.	6	8	<u>10</u>	8	<u>10</u>	.	.	6	<u>7</u>
Faye	5	6	<u>12</u>	.	.	.	.	.	<u>10</u>	.	.	3	<u>11</u>	.	.	<u>10</u>	.
Lamar	6	7	<u>12</u>	.	.	.	.	.	9	<u>10</u>	.	5	<u>10</u>	.	.	<u>11</u>	.
Lucas	3	7	9	<u>12</u>	.	.	.	.	<u>11</u>	.	.	9	7	9	<u>11</u>	<u>11</u>	.
Shirley	4	2	2	6	3	7	<u>11</u>	.	<u>10</u>	.	.	5	7	<u>10</u>	.	<u>11</u>	.

Underlined scores indicate that the subject passed criterion in that session. Criterion was at least 10/12 within one session or at least 8/12 across two consecutive sessions.

Table 3

First-trial performance in the unseen tracking and avoid-stimulus control tests

<i>Exp. 1: Lens test</i>	First unseen tracking trial	First avoid-stimulus trial
Artemus	0	0
Faye	0	1
Lamar	1	1
Lucas	1	1
Shirley	1	1
<i>Exp. 2: Mirror test</i>		
Mercury	1	.
Panzee	1	1
Sherman	1	.
<i>Exp. 3: Color test</i>		
Lamar	1	.
Lana	0	.
Lucas	1	0
Mercury	1	.
Sherman	1	.

1 = correct (chooses truly larger food reward); 0 = incorrect

Mirror test – number of trials (out of 12) in which chimpanzees chose the truly larger number of grapes, across all experimental sessions (S)

Table 4

Subject	Basic mirror test								Seen tracking		Unseen tracking			Avoid-mirror	
	S1	S2	S3	S4	S5	S6	S7	S8	S1	S2	S1	S2	S3	S4	S1
Faye	5	6	4	6	5	6	6	6	.	.	.	.	.	.	.
Lana	7	2	5	4	4	6	8	7	.	.	.	.	.	.	.
Lucas	9	6	3	8	4	1	3	4	.	.	.	.	.	.	.
Mercury	3	6	7	8	5	<u>10</u>	.	.	<u>10</u>	.	9	6	6	5	.
Panzee	3	<u>12</u>	.	.	.	.	.	.	8	<u>12</u>	<u>10</u>	.	.	.	<u>12</u>
Sherman	2	5	6	9	<u>11</u>	.	.	.	<u>12</u>	.	8	7	8	3	.

Underlined scores indicate that the subject passed criterion in that session. Criterion was at least 10/12 within one session or at least 8/12 across two consecutive sessions.

Table 5

Color test – number of trials (out of 8) in which chimpanzees chose the truly yellow (Y) or truly green (G) box across all sessions (S).

Basic color test										Seen tracking				Unseen tracking				Avoid-filter																
														unstacked				stacked																
	S1	S2	S3	S4	S5	S6	S7	S8		S1	S2	S1	S2	S3	S4	S1	S2	S3	S4															
Subject	Y	G	Y	G	Y	G	Y	G	Y	Y	G	Y	G	Y	G	Y	G	Y	G	Y	G	Y	G	Y	G									
Cissie	0	4	4	5	6	4	7	5	1	2	3	4	5	5	4	4	.	.	.	.	.	.	.	.	.									
Faye	0	8	5	7	4	8	5	4	5	8	4	8	3	8	4	8	.	.	.	.	.	.	.	.	.									
Lamar*	3	6	4	8	6	8	5	8	7	7	.	.	.	.	.	6	8	7	8	8	1	2	8	6	8	6	8	.	.	.	.	.		
Lana	7	8	.	.	.	.	.	.	.	.	.	.	.	.	.	8	6	6	6	8	5	4	8	4	7	6	5	.	.	.	.	.		
Lucas	5	6	5	8	8	8	.	.	.	.	.	.	.	.	.	8	8	.	.	8	6	8	8	.	.	.	7	6	7	8	8	3	8	7
Mercury	5	8	5	7	4	8	8	8	.	.	.	.	.	.	.	8	8	.	.	3	7	3	7	4	8	4	5	.	.	.	.	.	.	
Sherman	5	7	5	8	8	7	.	.	.	.	.	.	.	.	.	6	8	7	6	8	6	8	2	8	4	8	5	.	.	.	.	.	.	

* Lamar passed the unseen tracking test but was unavailable to participate in the avoid-filter test.
Underlined scores indicate that the subject passed criterion in that session (for both colors concurrently). Criterion was at least 7/8 within one session or at least 6/8 across two consecutive sessions.