Anim Cogn (2003) 6 : 77–86
DOI 10.1007/s10071-003-0159-x

ORIGINAL ARTICLE

Jennifer Vonk

Gorilla (*Gorilla gorilla gorilla*) and orangutan (*Pongo abelii*) understanding of first- and second-order relations

Received: 7 June 2002 / Revised: 10 November 2002 / Accepted: 12 January 2003 / Published online: 1 March 2003

Abstract Four orangutans and one gorilla matched images in a delayed matching-to-sample (DMTS) task based on the relationship between items depicted in those images, thus demonstrating understanding of both first- and second-order relations. Subjects matched items on the basis of identity, color, or shape (first-order relations, experiment 1) or same shape, same color *between* items (second-order relations, experiment 2). Four of the five subjects performed above chance on the second-order relations DMTS task within the first block of five sessions. High levels of performance on this task did not result from reliance on perceptual feature matching and thus indicate the capability for abstract relational concepts in two species of great ape.

Keywords Orangutan · Gorilla · Relational concepts

Introduction

Researchers investigating natural concept formation in non-human species have had difficulty disentangling the role of conceptual versus perceptual processing. Because the concept of how environmental stimuli are related is independent of their physical properties, some of the strongest evidence for "abstract" concept formation comes from researchers examining the extent to which non-human species can understand the relationships between items. An understanding of first-order relations is demonstrated when one perceives the relations between items. For instance, a mother and child may be perceived as exemplifying the "parent" relationship. An understanding of second-order relations involves understanding the relationship between two other relations and demonstrates an even more complex cognitive skill (Tomasello and Call 1997; Thompson and Oden 2000). One type of second-order relation is the ability to perceive analogies such that the relationship between two objects can be equivalent to the relationship between two other objects. For example, one understands that the relationship between a mother and her child is analogous to the relationship between a father and his child; both relations are that of parent and child and are understood to be equivalent.

Acquisition of first-order relational concepts such as identity (sameness) and oddity (difference) has been reliably demonstrated in monkeys through the use of either same/different (Bhatt and Wright 1992; Fujita 1983) or matching-to-sample (MTS) tasks (Burdyn and Thomas 1984; D'Amato et al. 1985, 1986; Thomas and Kerr 1976). In the former task, subjects are trained to touch two pictures on a touch screen and then to make a response to indicate whether the pictures were the same as or different from one another. Subjects select either symbols or response buttons that represent the concepts of "same" or "different." In the MTS task, a sample image is presented either simultaneously or successively (as in DMTS tasks) with two comparison images, a correct and an incorrect match, and subjects make a choice between the two comparison images. In both types of tasks, the generality of the concept is inferred from the subject's ability to perform accurately with novel stimuli in transfer tests. However, if the discrimination is perceptually based, it is possible for the monkeys to perform accurately in these experiments by way of simpler processes such as feature matching or entropy detection, without necessarily having acquired concepts for "sameness" or "difference."

Evidence for identity and oddity concepts in non-primate species has been even more equivocal. Initially it was concluded that pigeons were unable to understand relations, but more recent studies indicate that this conclusion may have been premature (Cook et al. 1995). It ap-

J. Vonk (✉)
York University,
4700 Keele Street, Toronto, ON, M3J 1P3, Canada

Present address:
J. Vonk
Cognitive Evolution Group, University of Louisiana at Lafayette,
4401 West Admiral Doyle Drive, New Iberia, LA 70560, USA
Tel.: +1-337-482026, Fax: +1-337-3730057,
e-mail: jxv9592@louisiana.edu

pears that when many items are presented in the stimulus array, pigeons are able to discriminate "same" from "different" displays (Cook et al. 1995, 1997; Wasserman et al. 1995). However, when the number of items in the stimulus array is reduced, neither pigeons nor baboons are able to do so (Young et al. 1997b; Wasserman et al. 2001b). Researchers have emphasized the role of entropy in that both species may attend to the amount of variance in the arrays (Young and Wasserman 1997; Young et al. 1997b; Wasserman et al. 2000, 2001a). The amount of variance differs between same and different displays and this difference is more pronounced when the stimulus array contains more items.

Alternatively, it may be that the perceptual regularity of a "same" stimulus array is emphasized when more items are included in the array, even when items are randomly arranged (Wasserman et al. 2000; Young and Wasserman 1997). The fact that non-human species might attend to either entropy or perceptual regularity argues against their use of a conceptual understanding of sameness and difference, and hence first-order relations. However, Young and colleagues have also presented pigeons with lists of successively presented "all same," "all different," or a mixture of same and different items. The pigeons were accurately able to indicate whether the list consisted of same or different items even though they could not respond until after presentation of the last item (Young et al. 1997a, 1999). This latter procedure provides more convincing evidence for non-perceptually based same/different conceptualization. In addition, Alex, an African grey parrot, responds accurately to such questions as "what is same?" between two objects (Pepperberg 1987), when the objects differ in terms of shape, color, or material. Thus it appears that some bird species might comprehend first-order relations.

Until very recently, evidence for second-order relational knowledge came from only chimpanzees (Gillan et al. 1981; Spinozzi 1993; Tanaka 1996; Thompson and Oden 1996, 2000; Thompson et al. 1997). Premack (1983) suggested that perhaps all vertebrates can easily perform MTS tests that require them to determine simply whether items look alike. Indeed, honeybees have been shown to understand sameness and difference in DMTS and non-MTS procedures, even when required to transfer these concepts across modalities (Giurfa et al. 2001). On the other hand, Premack felt that only language-trained chimpanzees should be able to solve second-order relational problems, such as analogies, because these latter concepts must be represented within an abstract code and rely upon conceptual versus perceptual matching.

Premack's hypothesis was based in part on the observation that one language-trained chimpanzee, Sarah, often succeeded at tasks which her non-language-trained counterparts failed. For example, when the chimpanzees were presented with a half-filled cylinder as the sample, and half an apple or three-quarters of an apple as alternatives, only Sarah was able to select consistently an alternative that matched the sample in terms of a proportion judgment (Premack 1983). When the apes were instead given, for example, one-quarter of an apple as a sample with one-quarter and one-half apples as alternatives, all of the chimpanzees passed the test. It was thus concluded that non-language-trained apes could make same/different judgments only on the basis of physical similarity, whereas the language-trained ape could make judgments on the basis of a representational code. Sarah was also able to solve analogical problems on the basis of both physical and relational properties (Gillan et al. 1981).

However, Premack's (1983) early contention that only language-trained chimpanzees code identity between abstract relations has been a controversial one. Oden et al. (1988) found that non-language-trained chimpanzees did not show transfer on a relational task, consistent with Premack's hypothesis. However, the same researchers (Oden et al. 1990) later determined that young chimpanzees *did* perceive similarities and differences between abstract relations, as evidenced by handling and looking times, despite failing to use this knowledge in an MTS task. Furthermore, Thompson et al. (1997) found that chimpanzees that had had experience with conditional and numeric token training could spontaneously match relations between relations, even when correct performance was not differentially rewarded. They could select correctly, for example, the BB versus the EF alternative when given AA as a sample, where these letters denote objects. Such conceptual matching problems could not be solved by attending solely to corresponding physical features, because the sample and alternative objects were not perceptually similar. Instead, the abstract relation between the objects must be represented for performance to be successful. The chimpanzees in this study were equally competent at matching relations and physical dimensions, despite having been consistently rewarded for correct matching of only the latter kind. The researchers concluded that what was needed was a token for "same" that provided the chimpanzee with a representational "tag" for the concept of the relation (see also Thompson and Oden 1996, 2000).

However, at least two studies provide evidence that the ability to form abstract relations is present in chimpanzees that have not received prior language or token training, or differential reinforcement. Tanaka (1996) found that one female chimpanzee was able to match items (and their photographs) on the basis of learned functional relationships, in an MTS task. For instance, she could match part one of a two-part object to its corresponding part or container, and she could match tools to their corresponding containers or assembled parts, even when non-matches more closely physically resembled the sample. Spinozzi (1993) found that some chimpanzees spontaneously sorted objects according to second-order classification schemes. For instance, when given six objects that differed in either one or two properties (color or form), they sometimes constructed two distinct class-consistent groupings.

In addition, two recent studies have suggested that monkeys, as well as apes, may be capable of a conceptual understanding of sameness. Bovet and Vauclair (2001) demonstrated that olive baboons were able to identify novel pairs of objects as "same" or "different" based on the objects' conceptual or functional relationships (food versus non-food). Judgments about membership in functional classes cannot be made merely on the basis of physical similarities between the items because category membership is defined by function, not appearance (Premack 1983). Fagot et al. (2001) provided support for the idea that baboons might also be sensitive to information about the relations between relations. In their task, baboons learned to match stimulus arrays to a sample based on whether the sample array included identical or different items. Because there was no overlap between individual stimulus icons in the sample and comparison arrays, accurate matching was deemed adequate evidence for the understanding of relations between icons in the arrays. However, the baboons required thousands of trials to reach criterion on the training set, so it appears that even if second-order relational concepts can be learned, it is only after significant training, and baboons may not spontaneously perceive such relations. In addition, in Fagot et al.'s (2001) experiment, "all same" arrays were perceptually distinct from "different" arrays.

In the current research, a touch screen was utilized to test the abilities of orangutans and a gorilla to match items on the basis of first- and second-order relations. Neither species has previously been tested for understanding of relational concepts. It was predicted that they would succeed on the first-order relations task after relatively few sessions, because choices could be based upon purely perceptual similarities. Accuracy could not be obtained on the second-order relations task by attending to physical cues because the incorrect comparison stimulus sometimes shared more perceptual features or entire compounds with the sample than did the correct comparison stimulus. Moreover, the sample and comparison stimuli always included only two items. As stated earlier, neither baboons nor pigeons can differentiate "same" from "different" displays with only two items in the stimulus array (Young, et al. 1997b; Wasserman et al. 2001b).

Experiment 1

This experiment examined whether orangutans and one gorilla could match items based on color or shape, in a delayed matching-to-sample (DMTS) procedure. Use of a delay between the sample and comparison stimuli made it necessary for subjects to maintain a representation of the sample (or an appropriate response) during the delay and thus made the task more difficult relative to standard simultaneous MTS tasks. The stimuli instantiated first-order relations such that the subjects had to perceive the relationship between the sample and the comparison stimuli. On one-third of the trials the correct comparison stimulus was an exact match for the sample. On another third of the trials the correct comparison stimulus matched the shape of the sample, and on the last third of the trials the correct comparison stimulus matched the color of the sample. On all trials, the incorrect comparison stimulus did not match the sample or the correct comparison in terms of either shape or color.

Methods

Subjects

Three adult male Sumatran orangutans (*Pongo abelii*), Dinding age 43, Molek, age 22, and Dinar, age 13; one adult female Sumatran orangutan, Abby, age 41; and one juvenile female western lowland gorilla (*Gorilla gorilla gorilla*), Zuri, age 4, served as subjects in this experiment. All were housed at the Toronto Zoo, Toronto, Ontario, Canada. Zuri was housed individually and the orangutans were also housed individually, when they were not on exhibit with four to six other orangutans during the day, 2 or 3 days a week.

Materials

Shapes of various colors were drawn using Deskpaint software for Macintosh. Four different shapes and colors each were used: circles, squares, triangles, and crosses, in blue, yellow, red, and green, for a total of 16 stimuli. Each object was approximately the same size, about 3.5×4.5 in. Identical shapes were exact copies of one another in different colors. The objects always filled the center of the image and were presented against a white background. An example of stimuli for one possible trial appears in Fig. 1. None of these images had been presented to any of the subjects previously. In fact the subjects had been exposed only to photographs of other animals, foods, and familiar objects and never to abstract drawings.

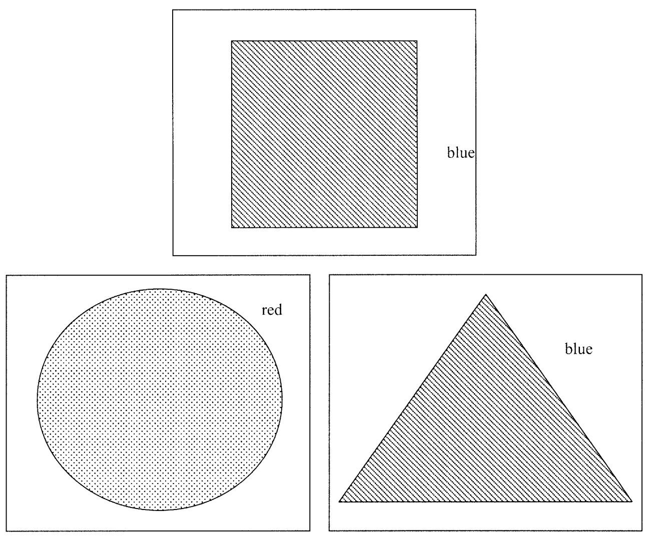

Fig. 1 Example of stimuli for one possible trial for experiment 1. The *blue square* is the sample stimulus. The *blue triangle* is the correct comparison stimulus and the *red circle* is the incorrect comparison stimulus. This is one possible color match trial. The comparison stimuli appear after a delay once the sample is selected and disappears

Table 1 Example of stimuli across trials for one possible session for experiment 1

Trial	Trial Type	Sample	Correct Comparison	Incorrect Comparison
1	Identity	Yellow triangle	Yellow triangle	Red cross
2	Color	Green circle	Green cross	Red square
3	Identity	Blue cross	Blue cross	Green circle
4	Shape	Green triangle	Yellow triangle	Blue square
5	Identity	Red circle	Red circle	Yellow square
6	Color	Blue triangle	Blue square	Yellow cross
7	Shape	Red square	Green square	Yellow circle
8	Shape	Yellow cross	Blue cross	Red triangle
9	Identity	Green square	Green square	Blue triangle
10	Shape	Blue circle	Red circle	Green cross
11	Color	Yellow square	Yellow circle	Green triangle
12	Color	Red cross	Red triangle	Blue circle

Procedure

The experiment was programmed in Filemaker Pro 3 software and was run on a Macintosh 5300 PowerBook computer. The images were presented on a 13-inch Apple touch screen monitor. The monitor was placed against the bars of the subjects' cages and the subjects were required either to reach through the mesh holes to touch the screen (the orangutans) or to reach underneath and touch the monitor (the gorilla). The experiment could be viewed on the laptop placed behind the monitor. The experimenter was positioned to one side of the screen and could not direct attention to either of the images on the screen.

All subjects had previously participated in experiments using the touch screen and had previously performed other DMTS tasks; therefore there were no training sessions (J. Vonk, unpublished data; Vonk and MacDonald 2002; unpublished data).

Each session consisted of 12 trials. During a trial, a sample image was presented in the center of the screen and stayed on the screen until the subject attended to the image and touched it, activating the touch screen. The two comparison images subsequently appeared on the screen after a brief delay (approximately 3 s). The subject was then required to select, by touching, only one of the two comparison images. If the subject selected the stimulus that matched the sample he or she was given a small food reward (M&M's candies or dried fruits and nuts for the orangutans, and dried fruits or nuts for the gorilla). After selection of the stimulus and after being rewarded, if correct, the screen advanced to the next sample image until all 12 trials were completed.

Within a session, each shape appeared as a sample stimulus three times and was paired with a different color each time. Each color was also presented in a sample image three times per session and paired with a different shape each time. There were three types of trials within a session. The correct comparison stimulus was either an exact match for the sample, or matched only the color or shape of the sample. The incorrect comparison stimulus always differed from both the sample and the correct comparison on both dimensions (shape and color). Therefore, within each session there were four trials in which only the color matched, four trials in which only the shape matched, and four trials in which the correct comparison stimulus was an exact match for the sample stimulus (identity). Twelve of the stimuli appeared twice each session, and the remaining four stimuli appeared three times each session. On each session, the stimuli were randomized so that the stimuli that appeared twice on one session appeared three times on other sessions. If a stimulus appeared three times it appeared in a different position each time so that it appeared as a sample, and on both the left and right side of the screen as test choices. If a stimulus appeared twice it might appear as a sample and as a comparison once or as a comparison twice on different trials on opposite sides of the screen. Order of presentation and pairings of stimuli were randomized on each trial. Presentation location was counterbalanced for left/right positions on the monitor. An example of trials within one *possible* session appears in Table 1.

Subjects were tested individually at the same time each day. The orangutans received 1–4 sessions per day, three or four times a week, and the gorilla completed between 5 and 12 sessions a day, 4 days a week. This discrepancy was due to the availability of the subjects and keepers and may have made it easier for the gorilla to acquire the task more quickly. Each subject was tested for a total of 30 (six blocks of 5) sessions, except for the female orangutan, Abby, who completed 25 (five blocks of 5) sessions. All subjects except for Abby completed experiment 1 before they began experiment 2. Abby was given 5 sessions of experiment 2 first and then sessions from each experiment were intermixed (typically 1 or 2 sessions of each experiment each day) so that experimental order effects would not be a factor in her performance.

Results and discussion

Table 2 presents the percentage of correct choices for each subject for each block of five sessions, both overall and for each type of discrimination. A binomial test for all sessions revealed that each subject performed at above-chance levels (50%), n=300 (for Abby) or 360 (for all other subjects), all P<0.001. Binomial tests were also con-

Table 2 Percent correct for each subject across blocks of five sessions, overall, and broken down by type of discrimination, experiment 1

Subject	Block	Overall	Identity	Shape	Color
Abby	1	89.96	95.0	100	75.0
	2	81.64	80.0	85.0	81.0
	3	74.98	70.0	85.0	70.0
	4	73.34	85.0	70.0	65.0
	5	69.98	71.6	64.4	74.0
Dinar	1	56.66	60.0	60.0	55.0
	2	65.02	70.0	70.0	55.0
	3	63.32	75.0	65.0	45.0
	4	51.68	55.0	50.0	50.0
	5	70.08	70.0	70.0	70.0
	6	56.70	48.0	55.0	70.0
Dinding	1	56.66	55.0	60.0	55.0
	2	55.00	70.0	51.6	43.0
	3	65.00	75.0	60.0	60.0
	4	57.70	65.0	50.0	58.4
	5	66.76	70.0	60.0	70.0
	6	71.70	75.0	62.0	78.4
Molek	1	74.98	75.0	80.0	70.0
	2	66.68	55.0	60.0	85.0
	3	56.68	75.0	56.3	50.0
	4	66.66	60.0	65.0	75.0
	5	73.34	85.0	65.0	70.0
	6	75.00	75.0	75.0	75.0
Zuri	1	62.88	75.0	45.0	65.0
	2	65.02	65.0	75.0	55.0
	3	63.32	60.0	70.0	60.0
	4	58.42	55.0	65.0	55.0
	5	78.30	75.0	85.0	75.0
	6	65.00	55.0	75.0	85.0

ducted to determine how many sessions were required for each subject to obtain above-chance levels of responding. Abby was above chance on her 1st session, n=12, P=0.006, Molek by the 3rd session, n=36, P=0.005, Zuri by the 6th session, n=72, P=0.04, Dinar by the 9th session, n=108, P=0.03, and Dinding by the 14th session, n=168, P=0.05. Abby averaged 78% correct across five blocks, compared to 69% for Molek, 65% for Zuri, 62% for Dinding, and 61% for Dinar. Independent samples t-tests comparing performance on the first block to performance on the last block of sessions revealed a significant increase in performance for Dinding, t_8=–2.50, P=0.04, and Dinar, t_8=–2.57, P=0.04, and a significant decrease for Abby, t_8=2.68, P=0.04. Thus, there was evidence of learning only for Dinding and Dinar. The decline in Abby's performance was likely due to poor health as she was beginning to show signs of weight loss and loss of appetite during the course of this experiment.

If purely perceptual processes dictated performance on the task one might expect the identity trials to be performed more accurately than the other types of trials, because the sample matched the comparison on two dimensions instead of one on only these trials. A series of repeated measures analyses of variance (ANOVAs) for each subject were conducted on the percent correct data for each session with type of trial (color match, shape match or identity) as factors. There was no effect of type of trial for any of the subjects, highest $F_{2,164}$=1.78, P>0.10. Overall, there were no significant differences between average percent correct across the types of trials, identity (68%), color (65%), or shape (67%) matches.

Because the incorrect comparison always differed from the sample on two dimensions whereas the correct comparison always matched the sample on at least one dimension, the correct choice was always perceptually more similar to the sample than was the incorrect choice. Thus the subjects would have been able to perform accurately on this task by making perceptual discriminations. Abby's high performance on the very first session suggested that first-order relations, or perceptual matches, could be perceived spontaneously and did not have to be trained. She averaged 90% on the first block. The other subjects were not as accurate initially. Average performance by all of the subjects taken together was only 67% across all sessions. Results of this experiment thus demonstrated individual differences. However, all subjects consistently performed the task at above chance levels, indicative of the ability to match stimuli on the basis of perceptual features such as same color and shape. They were subsequently tested for their ability to understand more complex second-order relations.

Experiment 2

This experiment examined the ability of the subjects to match stimuli according to the relationship between items. Subjects were required to match stimuli according to whether the two objects within the sample stimulus were the same shape or the same color as each other. On some trials the incorrect comparison consisted of the same shapes or colors as the sample image while only the relationship *between* the objects in the images conveyed the relevant information for choosing the correct comparison. For example, consider the stimuli in Fig. 2. The sample in this example consists of a blue and yellow square, two images of the same shape, but different colors. The correct comparison stimulus is the red and green circle pair since they too are the same shape as each other. However the exact shape itself is different from that in the sample (circles vs squares). Furthermore, the incorrect comparison stimulus consists of a blue triangle and a blue circle, thus exemplifying a same-color versus a same-shape relationship. This incorrect comparison stimulus does, however, include blue images, which might be perceived as perceptually similar to the blue square in the sample. Therefore this task was thought to be more cognitively demanding than purely perceptual matching because correct performance was dependent on processing the relationship between the images and not the physical features of particular images.

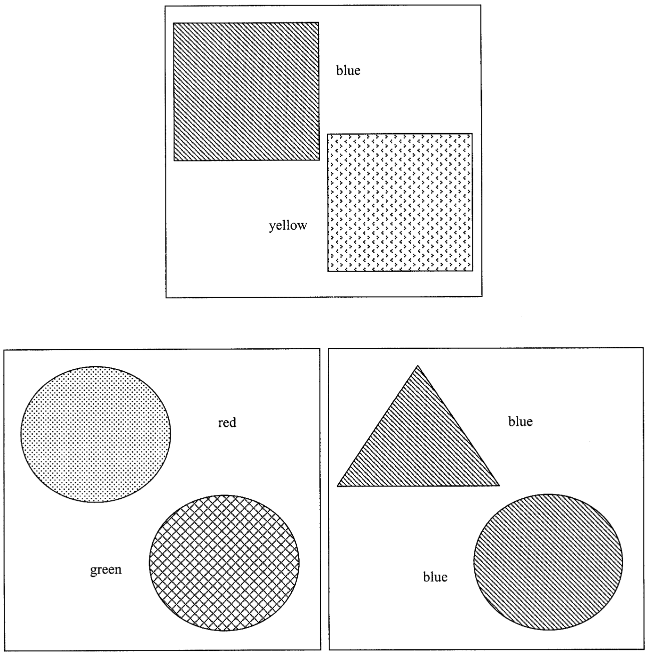

Fig. 2 Example of stimuli for one possible trial for experiment 2. The two *squares* constitute the sample stimulus. The two *circles* constitute the correct comparison stimulus and the two *blue shapes* constitute the incorrect comparison stimulus. The comparison stimuli appear after a delay once the sample is selected and disappears. The relation being tested is that of shape, and the incorrect stimulus represents the color match. Other trials include incorrect stimuli that do not match on either shape or color

Methods

Subjects

The same five subjects participated in experiment 2.

Materials

As in experiment 1, stimuli were created using Deskpaint software. Each stimulus consisted of a pair of shapes against a white background. All of the shapes were of uniform size (approximately 2.5×3 in.) and appeared in one of four colors: red, blue, green, or yellow. Three different shapes were used: squares, circles, and triangles. The stimulus set of 18 images included 6 pairs of images that were the same shape but different colors (e.g., a red and a blue circle), 6 pairs of images that were the same color but different shapes (e.g., a blue circle and a blue square), and 6 pairs of images that matched on neither shape nor color (e.g., red circle and green triangle). The latter 6 image types were presented only once each session. Of the remaining 12 images, 3 of each type were presented three times, and the other 3 of each type were presented twice. The images presented three times appeared as a sample, a correct comparison stimulus, and as an incorrect stimulus during the session. The images presented twice appeared as samples and as correct stimuli. Because the stimuli were sorted randomly for each session, the items that appeared twice during one session appeared three times during other sessions. An example of stimuli for one *possible* trial appears in Fig. 2. This figure represents only one possible combination of stimuli among many. On some trials, whole compounds were shared between stimuli, while on other trials, each compound of the three stimuli was unique. Because of the randomization process, the distribution of trials with varying degrees of perceptual overlap varied across sessions. For an example of trials within one *possible* session, see Table 3.

Table 3 Example of stimuli across trials for one possible session for experiment 2. *B* blue, *R* red, *G* green, *Y* yellow, *c* circle, *t* triangle, *s* square

Trial	Sample	Correct comparison	Incorrect comparison
1	(Bc) and (Bt)	(Bt) and (Bs)	(Bs) and (Gs)
2	(Rt) and (Yt)	(Rs) and (Gs)	(Rt) and (Rs)
3	(Rc) and (Rs)	(Rt) and (Rs)	(Rc) and (Bc)
4	(Bc) and (Yc)	(Rc) and (Bc)	(Yc) and (Yt)
5	(Gc) and (Gs)	(Yc) and (Yt)	(Rs) and (Gs)
6	(Bs) and (Gs)	(Bt) and (Yt)	(Bt) and (Yc)
7	(Rc) and (Bc)	(Bc) and (Yc)	(Bs) and (Yc)
8	(Yc) and (Yt)	(Gc) and (Gs)	(Rs) and (Yt)
9	(Bt) and (Gt)	(Bs) and (Gs)	(Rc) and (Rs)
10	(Rs) and (Gs)	(Bt) and (Yt)	(Gc) and (Bt)
11	(Rt) and (Rs)	(Gs) and (Gc)	(Rs) and (Bc)
12	(Bt) and (Bs)	(Bc) and (Bt)	(Gt) and (Rc)

Procedure

This experiment followed the same procedure as described for experiment 1, except for the difference in materials noted above. During each session, half of the 12 samples were images of two identical shapes of different colors (shape match), and the other half were images of two different shapes in the same color (color match). For each of these two types of trials, half of the incorrect stimuli consisted of two items that matched on the opposite variable to that matched within the sample stimulus. For instance, on color match trials, half of the incorrect stimuli were pairs of items that were the same shape but different colors. On the other six trials, the incorrect stimuli consisted of pairs of items that differed in both shape and color. Each subject completed five blocks of 5 sessions for a total of 25 sessions, except for Abby who completed four blocks of 5 sessions for a total of 20 sessions.

Results and discussion

As can be seen in Table 4, and confirmed with binomial tests, each subject, except for Dinar, n=300, P>0.10, performed at levels significantly above chance, n=300 for Molek, Dinding, and Zuri, n=240 for Abby, all P<0.001, across sessions. Separate binomial tests were also con-

Table 4 Percent correct for each subject across blocks of five sessions, overall, and by discrimination type

Subject	Block	Overall	Shape match	Color match
Abby	1	78.3	76.0	76.6
	2	73.3	80.0	66.8
	3	71.7	70.0	73.2
	4	73.3	66.6	80.0
Dinar	1	53.3	53.4	56.0
	2	56.7	63.4	50.0
	3	53.3	50.0	56.8
	4	55.0	53.4	56.6
	5	55.0	50.0	60.0
Dinding	1	63.3	73.2	53.4
	2	56.7	73.4	40.0
	3	66.7	61.0	70.2
	4	65.0	80.0	49.2
	5	70.8	76.8	63.2
Molek	1	70.0	63.4	76.6
	2	85.0	80.0	90.0
	3	76.7	73.4	76.6
	4	73.3	72.6	71.8
	5	80.0	76.8	83.2
Zuri	1	65.0	67.8	63.2
	2	75.0	77.6	73.4
	3	76.7	79.4	76.6
	4	75.0	70.0	79.8
	5	69.4	73.2	66.8

ducted for each subject to determine how many sessions were required before performance reached above-chance levels. Abby's performance was significantly above chance by the second session, n=24, P=0.02, Molek's by the third session, n=36, P=0.03, Zuri's by the fourth session, n=47, P=0.04, and Dinding's by the fifth session, n=60, P=0.05. Thus on this task, four of the five subjects reached above-chance levels within the first block of five sessions. Independent samples t-tests comparing performance on the first block to performance on the last block of sessions revealed no significant increase in performance for any of the subjects, highest t_8=–2.34, P=0.06, suggesting that accuracy was not due to learning.

Each subject's performance as a function of the type of matching required, across blocks of five sessions, is shown in Table 4. A repeated measures ANOVA of subjects' scores on each session by type of sample (color or shape match) was conducted separately for each subject. The effect of sample type was significant only for Dinding, $F_{1,24}$=10.62, P=0.003, and for Molek, $F_{1,24}$=5.88, P= 0.02. Dinding performed at 74% correct on shape match trials, but at only 55% correct on color match trials. Molek performed at 73% correct on shape match trials and at 80% correct on color match trials.

A repeated measures ANOVA of subjects' percent correct scores, with type of incorrect stimulus (other match, no match) as factors, was also conducted for each subject. None of the subjects' performances were significantly affected by the type of incorrect stimulus, highest $F_{1,24}$=3.24, P=0.09.

Table 5 Percent correct for each subject as a function of the number of matching compounds between the sample and the correct and incorrect comparisons

Subject	Number of matching compounds between sample and correct comparison		Number of matching compounds between sample and incorrect comparison	
	0	1	0	1
Abby	73.3	82.6	79.9	76.1
Dinar	53.8	62.6	53.6	62.8
Dinding	63.9	73.9	63.6	74.2
Molek	76.0	81.3	78.7	78.6
Zuri	71.4	82.6	70.0	84.1

On some trials, one of the two compounds in the sample also appeared in one or both of the correct and incorrect comparison stimuli. For instance, if the sample consisted of a *blue square* and a *blue triangle*, the correct comparison might consist of a *blue square* and a blue circle, and the incorrect comparison might consist of a *blue triangle* and a green circle; thus each comparison stimulus would share one compound with the sample. This is just one example of several possible stimulus combinations. On some trials, none of the compounds were shared across the three stimuli and on other trials a compound was shared only between the sample and *either* the correct or the incorrect comparison stimulus, but not both. (See Table 2 for one sample session.)

Table 5 presents each subject's percentage of correct choices depending upon the number of shared compounds between the sample and each of the comparison stimuli. An ANOVA of scores on each trial with compound matches (one or none) between the sample and the correct comparison and with compound matches between the sample and the incorrect comparison as independent variables was conducted for each subject. These analyses revealed no significant effects for any of the orangutan subjects, highest $F_{1,308}$=1.44, P=0.23. Thus they were no more likely to choose the correct comparison if it shared a physical compound of the sample stimulus than if it did not, and they were no more likely to avoid selecting the incorrect comparison if it did not share a compound with the sample. Although the gorilla's responses were affected by an interaction between the presence of a shared compound between the sample and the correct comparison, and between the sample and the incorrect comparison, $F_{1,308}$=3.78, P=0.05, Bonferroni post hoc tests revealed no significant differences between the conditions.

In addition to an entire compound matching between the various stimuli, several independent dimensions could also overlap. For instance, the color or shape of one or both compounds of the sample could either match or not match the color or shape of one or both compounds of each of the comparison stimuli on a given trial. For each comparison between the stimuli (e.g., sample–correct comparison, sample–incorrect comparison, and correct–incorrect comparison) the number of possible matching

Table 6 Percent correct for each subject as a function of the number of matching dimensions between the sample and the correct and incorrect comparisons

Subject	Number of matching dimensions between sample and correct comparison			Number of matching dimensions between sample and incorrect comparison			
	0	1	2	0	1	2	3
Abby	76.7	71.0	82.3	88.2	71.6	73.4	73.3
Dinar	43.5	54.9	56.5	62.8	39.4	59.7	44.4
Dinding	70.0	69.9	72.0	76.9	63.0	58.2	91.7
Molek	61.6	73.7	80.8	82.2	79.0	75.1	51.9
Zuri	60.3	79.3	66.0	85.9	72.8	53.6	61.9

dimensions was determined. For instance, if the same color appeared in two stimuli, this was scored as one matching dimension, and likewise for shape. Table 6 presents each subject's percentage of correct choices as a function of the number of dimensions shared between the sample and each of the two comparison stimuli.

Univariate ANOVAs were conducted for each subject individually with score (correct or incorrect) on each trial as the dependent measure and number of matching dimensions between the sample and each of the correct and incorrect comparisons as separate independent variables. For Abby, Molek, Dinding, and Zuri, the number of shared dimensions between the sample and each of the comparisons did not significantly affect performance, highest F=2.34, P=0.07. Thus they were no more accurate when the sample shared more perceptual features with the correct comparison, or when the sample shared less perceptual features with the incorrect comparison. Dinar's performance was affected by an interaction between the number of shared dimensions between the sample and the correct comparison and the sample and the incorrect comparison, $F_{6,288}$=2.31, P=0.03. There was no obvious logical pattern in Dinar's responses according to these interactions, making them difficult to interpret. Because his responding was at chance, it was possible that these interactions reflected random biases. In general, it appeared that the subjects' responses were not controlled by the number of matching perceptual dimensions between various combinations of the stimuli on any given trial.

To determine if the subjects' performance differed between experiments, an ANOVA compared their average percent correct scores for each session across experiments. For Dinding and Abby there was no difference between performances on experiments 1 and 2, $F_{1,43}$=1.74 and 1.24, respectively, both P>0.05. Dinding averaged 60 versus 64% correct and Abby averaged 78 versus 74% correct on experiments 1 and 2, respectively. Zuri and Molek had higher average scores for experiment 2, $F_{1,43}$=4.85 and 5.69, respectively, P=0.03 and 0.02. (for Zuri, 66 and 73% correct, for Molek 68 and 76% correct). Dinar had higher average scores for experiment 1, $F_{1,43}$= 6.49, P=0.01 (61 vs 55% correct).

The fact that the subjects' performance, except for that of Dinar, was consistently above chance within the first block of sessions indicates that they perceived the second-order relations and did not have to learn to associate exemplars through extended training. The individual differences might suggest that although some orangutans are capable of perceiving second-order relations, these discriminations are not of critical importance for this species and are not always made spontaneously. However, Molek, Abby, and Zuri made the discriminations at fairly high levels initially, without any training on the task. Furthermore, performance for two of these subjects was, on average, higher than their performance on the first-order-relations task in experiment 1. It is possible that this finding was due to transfer across experiments and previous experience with similar stimuli. As a control for this possibility Abby participated in experiments 1 and 2 simultaneously and notably there was no significant difference between her performances across the two experiments.

General discussion

The four orangutans and one gorilla tested in the current experiments demonstrated some understanding of both first- and second-order relations. In experiment 1, they were able to match items based on a relationship between the sample and comparison stimuli (same shape, color, or identity) at above-chance levels. However, it is possible that these matches were made by matching perceptual features and did not necessarily require the understanding of a "sameness" concept. In experiment 2, the subjects were able to perceive the relationship between items in the sample stimulus pair and match this relation to that between items in the comparison stimulus pair. This latter finding demonstrates that orangutans and gorillas may be capable of abstracting the relation between relations and are not reliant upon purely perceptual processing when classifying stimuli. Instead, they seemed able to apply the concept of "same" even when incorrect stimuli more closely resembled the sample stimuli perceptually. The number of shared physical compounds or dimensions between the sample and the comparisons did not significantly affect the subjects' responses.

These experiments controlled for some of the problems inherent in natural concept images. With natural stimuli it can be difficult to de-couple the role of perceptual and conceptual processing. In such experiments, it is also difficult to determine exactly which features the subjects are using to discriminate among the stimuli. In the current study, the number of features available for analysis was limited, and accurate performance could only be obtained by attending to the relevant variables. In the second experiment, the relevant variable was an abstract relation and successful performance required an understanding of the concepts "same," "color," and "shape." In the first experiment, subjects needed to attend only to which stimuli were perceptually more similar. Surprisingly, two of the subjects performed better on the purportedly more difficult conceptual task. These two animals succeeded on the second-order relations matching task within the first few sessions and therefore did not have to learn the concept being tested. It is possible that prior experience with similar stimuli in experiment 1 led to some degree of non-specific transfer across experiments. Abby's data, however, demonstrated that this experience was not necessary because she was able to perform equally well on the second-order relations task even though she completed the first five sessions of this task prior to participating in experiment 1.

Interestingly, a similar study was conducted with an African grey parrot, Alex, who has undergone extensive language training (Pepperberg 1987). In this experiment, the parrot was shown two objects that differed in shape but were the same color on some trials, and two objects that differed in color but were the same shape, on other trials. He was asked, "What is same?" and responded appropriately, even to novel pairs of items. Thus, Alex demonstrated understanding of the concepts "same," "different," "shape," and "color," as well. However, Alex was not required to produce or match pairs of objects on the basis of the relationship between items (Roitblat and Von Fersen 1992), as were the subjects in the current study.

Premack and Premack (1983) argued that only language-trained apes could use second-order relations in analogy tasks. They tested the ability of language-trained chimpanzees to match, for example, a pair of oranges to a pair of apples when the incorrect stimulus was an apple and a banana. The chimpanzees were also given the reverse task of matching in which the sample stimulus was an orange and an apple, the correct stimulus was a pear and a pineapple, and the incorrect stimulus was a pair of pears, for example. These chimps were able to succeed at both tasks. This experiment showed that the chimpanzees understood the concepts "same" and "different" and could apply these concepts to novel stimuli that did not match the sample stimuli on a perceptual basis. Here, subjects were not tested for their understanding of "different" as well as "same."

Recently Fagot et al. (2001) demonstrated that baboons could match images on the basis of relations between relations. The two baboon subjects were able to match sample stimuli that included a number of all-same or all-different items to a comparison stimulus in which items shared the same relationship (all same or all different), even though there was no overlap between items within a stimulus array. They were able to transfer to novel items, but their performance declined when the number of items within a stimulus array was reduced. This finding might be due to a reduction in perceptual regularity. When many items are displayed in the stimulus array, stimulus arrays that include many identical items look very distinct from stimulus arrays that include many different items. This distinctiveness is reduced when only two same or different items are presented, as in the current study. This study extends this finding further by showing that orangutans and a gorilla could also match stimuli according to *what* dimension was shared between objects pictured in the same stimulus, which goes beyond simply distinguishing two identical objects from two non-identical objects.

In addition, the baboons in Fagot et al.'s (2001) study required thousands of trials to reach criterion on the task even though they had had prior experience with similar stimuli and with making same–different judgements. The subjects in the current experiments had no prior experience with the stimuli used here and did not require training to reach levels well above chance. Furthermore, the baboons underwent many more trials a day and on successive days whereas the orangutans in particular were sometimes given only one or two sessions a day, with more than a week between sessions in some cases. At the very least, if both baboons and great ape species can be said to understand second-order relations, this ability seems to be utilized more readily by the apes. It is also notable that the gorilla in this study was only 4 years old at the time of testing. Studies of object sorting have indicated that chimpanzees do not sort according to second-order relations until at least the age of 5 (Spinozzi et al. 1999).

Previous research has shown that, when given the opportunity to sort objects according to various dimensions, non-human primates tend to sort more often according to first-order relative to second-order relations (McClure and Culbertson 1977; Garcha and Ettlinger 1979; Spinozzi and Natale 1989). However, chimpanzees have shown some evidence for spontaneous sorting along second-order relations (Spinozzi 1993). Tanaka (1995, 1996) has also shown that chimpanzees may be able to match items on the basis of complementarity and previously learned relations. Where great apes may distinguish themselves from monkeys may be in the ability to understand and make use of second-order relations spontaneously. However, other monkey species have not been adequately tested with such tasks (Tomasello and Call 1997). Furthermore, Premack's contention has never been adequately tested in other species. No monkeys that have been human reared and/or language trained have ever been tested in discrimination paradigms.

The results of the current experiments suggest that non-language-trained apes of two previously untested species can succeed at conceptual discrimination tasks. Contrary to Premack's (1983; Premack and Premack 1983; Thompson and Oden 1996, 2000) hypothesis, these experiments suggest that language or symbolic token training is *not* a necessary condition for the acquisition and instrumental use of concepts pertaining to the relationships between items. The disparity between this conclusion and that of some researchers working with chimpanzees (Thompson and Oden 1996, 2000; Thompson et al. 1997) may be due to species differences in conceptual abilities, but, alternatively and possibly more likely, this disparity may be due to differences in the experimental procedures. In the current experiments, subjects were not required to produce or complete analogies from a selection of several possible alternatives but instead were required to select one of only two stimuli whose relationship matched that of the sample stimulus. Perhaps this constraint on the number of possible alternatives eliminated the need for a symbolic representational system.

Alternatively, conclusions regarding the role of abstract symbol training in mediating analogical abilities in our closest relatives may have been premature, especially given that these conclusions were based upon a relatively small sample size (e.g., one "symbol naive" chimpanzee in Thompson et al. 1997). Furthermore, several recent studies indicate that monkeys, as well as non-language-trained chimpanzees, may have greater capacities for making conceptual judgments than was previously believed (Spinozzi 1993; Tanaka 1996; Bovet and Vauclair 2001; Fagot et al. 2001).

Acknowledgements This research was supported by a Natural Sciences and Engineering Research Council (NSERC) postgraduate scholarship granted to the author. The cooperation and support of the staff at the Toronto Zoo was greatly appreciated. Special thanks to Dianne Devison, Bev Carter, Vanessa Phelan, Gerri Mintha, Charles Guthrie, Tim McCaskie, Connie Wiederman, Heidi Minicke, Michelle Smith, Karyn Tunwell, Bridget Burke-Johnson, Andrea Beatson, Rick Vos, David Partington, Des Macguire, John Armstrong, and Jackie Craig, without whose assistance this study would not have been possible. I am also grateful to Suzanne MacDonald for the opportunity to undertake this project, and for providing the equipment and her guidance. Thanks to Robert Sorge, Roger Thompson, and several anonymous reviewers for comments on earlier drafts. These experiments conform to the Canadian laws and regulations concerning the use of animals.

References

Bhatt RS, Wright AA (1992) Concept learning by monkeys with video picture images and a touch screen. J Exp Anal Behav 57:219–225

Bovet D, Vauclair J (2001) Judgement of conceptual identity in monkeys. Psychon Bull Rev 8:470–475

Burdyn LE Jr, Thomas RK (1984) Conditional discrimination with conceptual simultaneous and successive cues in the squirrel monkey (*Saimiri sciureus*). J Comp Psychol 4:405–413

Cook BR, Cavoto KK, Cavoto BR (1995) Same-different texture discrimination and concept learning by pigeons. J Exp Psychol Anim Behav Process 21:253–260

Cook RG, Katz JS, Cavoto BR (1997) Pigeon same-different concept learning with multiple stimulus classes. J Exp Psychol Anim Behav Process 23:417–433

D'Amato MR, Salmon DP, Colombo M (1985) Extent and limits of the matching concept in monkeys (*Cebus apella*). J Exp Psychol Anim Behav Process 11:35–51

D'Amato MR, Salmon DP, Loukas E, Tomie A (1986) Processing of identity and conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*). Anim Learn Behav 14:365–373

Fagot J, Wasserman EA, Young ME (2001) Discriminating the relation between relations: the role of entropy in abstract conceptualization by baboons (*Papio papio*) and humans (*Homo sapiens*). J Exp Psychol Anim Behav Process 27:316–328

Fujita K (1983) Formation of the sameness-difference concept by Japanese monkeys from a small number of color stimuli. J Exp Anal Behav 40:289–300

Garcha HS, Ettlinger G (1979) Object sorting by chimpanzees and monkeys. Cortex 15:213–224

Gillan DJ, Premack D, Woodruff G (1981) Reasoning in the chimpanzee 1. Analogical reasoning. J Exp Psychol Anim Behav Process 7:1–17

Giurfa M, Zhang S, Jenett A, Menzel R, Srinivasan MV (2001) The concepts of 'sameness' and 'difference' in an insect. Nature 410:930–933

McClure MK, Culbertson G (1977) Chimpanzees' spontaneous temporal sorting of stimuli on the basis of physical identity. Primates 18:709–711

Oden DL, Thompson RKR, Premack D (1988) Spontaneous transfer of matching by infant chimpanzees. J Exp Psychol Anim Behav Process 14:140–145

Oden DL, Thompson RKR, Premack D (1990) Infant chimpanzees spontaneously perceive both concrete and abstract same/different relations. Child Dev 61:621–631

Pepperberg I (1987) Acquisition of the same/different concept by and African grey parrot (*Psittacus erithacus*): learning with respect to categories of color, shape, and material. Anim Learn Behav 15:423–432

Premack D (1983) The codes of man and beasts. Behav Brain Sci 6:125–167

Premack D, Premack AJ (1983) The minds of an ape. Norton, New York

Roitblat HL, Von Fersen L (1992) Comparative cognition: representations and processes in learning and memory. Annu Rev Psychol 43:671–710

Spinozzi G (1993) Development of spontaneous classificatory behavior in chimpanzees (*Pan troglodytes*). J Comp Psychol 107: 193–200

Spinozzi G, Natale F (1989) Classification. In: Antinucci F (ed) Cognitive structure and development in nonhuman primates. Erlbaum, Hillsdale, N.J., pp 163–187

Spinozzi G, Natale F, Langer J, Brakke KE (1999) Spontaneous class grouping behavior by bonobos (*Pan paniscus*) and common chimpanzees (*P. troglodytes*). Anim Cogn 2:157–170

Tanaka M (1995) Object sorting in chimpanzees (*Pan troglodytes*): classification based on physical identity, complementarity, and familiarity. J Comp Psychol 109:151–161

Tanaka M (1996) Information integration about object-object relationships by chimpanzees (*Pan troglodytes*). J Comp Psychol 110:323–335

Thomas RK, Kerr S (1976) Conceptual conditional discrimination in *Saimiri sciureus*. Anim Learn Behav 4:333–336

Thompson RKR, Oden DL (1996) A profound disparity revisited: perception and judgement of abstract identity relations by chimpanzees, human infants and monkeys. Behav Process 35: 149–161

Thompson RKR, Oden DL (2000) Categorical perception and conceptual judgements by nonhuman primates: the paleological monkey and the analogical ape. Cogn Sci 24:363–396

Thompson RKR, Oden DL, Boysen ST (1997) Language-naïve chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. J Exp Psychol Anim Behav Process 23:31–43

Tomasello M, Call J (1997) Primate cognition. Oxford University Press, Oxford

Vonk J, MacDonald SE (2002) Natural concept formation in a juvenile gorilla (*Gorilla gorilla gorilla*) at 3 levels of abstraction. J Exp Anal Behav 78:315–332

Wasserman EA, Hugart JA, Kirkpatrick-Steger K (1995) Pigeons show same-different conceptualization after training with complex visual stimuli. J Exp Psychol Anim Behav Process 21: 248–252

Wasserman EA, Young ME, Nolan BC (2000) Display variability and spatial organization as contributors to the pigeon's discrimination of complex visual stimuli. J Exp Psychol Anim Behav Process 26:133–143

Wasserman EA, Fagot J, Young ME (2001a) Same-different conceptualization by baboons (*Papio papio*): the role of entropy. J Comp Psychol 115:42–52

Wasserman EA, Young ME, Fagot J (2001b) Effects of number of items on the baboon's discrimination of same from different visual displays. Anim Cogn 4:163–170

Young ME, Wasserman EA (1997) Entropy detection by pigeons: response to mixed visual displays after same-different discrimination training. J Exp Psychol Anim Behav Process 23:157–170

Young ME, Wasserman EA, Dalrymple R (1997a) Memory-based same-different conceptualization by pigeons. Psychon Bull Rev 4:552–558

Young ME, Wasserman EA, Garner KL (1997b) Effects of number of items on the pigeon's discrimination of same from different visual displays. J Exp Psychol Anim Behav Process 23: 491–501

Young ME, Wasserman EA, Hilfers MA, Dalrymple R (1999) The pigeon's variability discrimination with lists of successively presented visual stimuli. J Exp Psychol Anim Behav Process 25:475–490