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

Petra H. J. M. Vlamings, Jana Uher, and Josep Call
Max Planck Institute for Evolutionary Anthropology

S. T. Boysen and G. G. Berntson (1995) found that chimpanzees performed poorly on a reversed contingency task in which they had to point to the smaller of 2 food quantities to acquire the larger quantity. The authors compared the performance of 4 great ape species (*Pan troglodytes*, *Pongo pygmaeus*, *Pan paniscus*, and *Gorilla gorilla*) on the reversed contingency task while manipulating food quantity (0–4 or 1–4) and food visibility (visible pairs or covered pairs). Results showed no systematic species differences but large individual differences. Some individuals of each species were able to solve the reversed contingency task. Both quantity and visibility of the food items had a significant effect on performance. Subjects performed better when the disparity between quantities was smaller and the quantities were not directly visible.

Keywords: inhibition, perseveration, problem solving

Finding food, manipulating fruits and plants, or competing to reproduce: To live is to solve problems. One of the skills required for problem solving is inhibition, which in the current study is defined as the ability to suppress certain prepotent responses in problem-solving situations (Deacon, 1997; Diamond, 1988; Santos, Ericson, & Hauser, 1999). Developmental studies on animals and children have shown that inhibitory problems can prevent subjects from solving certain tasks (Deacon, 1997; Diamond, 1990). A paradigm that is used to study inhibitory skills is the reversed contingency task, which has been investigated in several nonhuman primate species, including chimpanzees (*Pan troglodytes*), Japanese macaques (*Macaca fuscata*), tamarins (*Saguinus oedipus*), squirrel monkeys (*Saimiri sciureus*), lemurs (*Eulemur fulvus* and *E. macaco*), and orangutans (*Pongo pygmaeus*; Anderson, Awazu, & Fujita, 2000; Boysen & Berntson, 1995; Boysen, Berntson, Hannan, & Cacioppo, 1996; Boysen, Berntson, & Mukobi, 2001; Boysen, Mukobi, & Berntson, 1999; Genty, Palmier, & Roeder, 2004; Kralik, Hauser, & Zimlicki, 2002; Shumaker, Palkovich, Beck, Guagnano, & Morowitz, 2001; Silberberg & Fujita, 1996).

Boysen and Berntson (1995) first ran an experiment in which 2 number-trained chimpanzees (*Pan troglodytes*) had to learn the following rule: Pick the smaller food quantity; receive the larger one. One chimpanzee acted as the selector and the other as the receiver. The selector was allowed to choose one of two different food quantities. As soon as the selector made its choice, the experimenter gave the content of the selected food array to the receiver, whereas the selector was given the content of the nonselected array. Therefore, it was to the selector's advantage to point to the food amount it did not want to have. Results showed that both chimpanzees kept picking the larger food quantity. This result persisted even in the absence of the competitor when the selected array was thrown back into the container with the reward supply (Boysen et al., 1996). Even after 400 trials, the chimpanzees kept choosing the larger food quantity. Three-year-old children as well as children with autism have also been reported to have significant problems inhibiting a prepotent response tendency to reach for a larger food quantity when presented with two food arrays of a different quantity (Russell, Mauthner, Sharpe, & Tidswell, 1991).

When the reversed contingency task was presented with different item sizes instead of different food amounts, item size appeared to exert a more powerful interference effect than item quantity, causing more severe inhibition problems (Boysen et al., 2001). However, when Arabic numerals were used instead of food arrays, performance increased significantly (Boysen & Berntson, 1995; Boysen et al., 1999). Boysen and Berntson (1995) suggested that the intrinsic incentive and/or perceptual features of the larger array made it difficult for subjects to inhibit choosing the larger array. Conversely, Arabic numerals represent numerosity symbolically and might increase performance by endorsing abstract processing of numerosity with less interference from the incentive properties of the stimuli (Boysen & Berntson, 1995).

Silberberg and Fujita (1996) tested Japanese macaques (*Macaca fuscata*) with the reversed contingency procedure analogous to the one used by Boysen and Berntson (1995). Instead of using arrays of one to six food items, they presented subjects with one unchanging quantity combination: four raisins versus one raisin. Like the chimpanzees, the macaques failed. In a follow-up condition,

Silberberg and Fujita changed the reward contingencies. Choosing the larger food quantity meant receiving no food at all. When the costs of choosing the larger food quantity were increased, subjects readily learned to pick the smaller food quantity. Silberberg and Fujita argued that the performance of the macaques improved because they were better able to control their responses due to the fact that the costs of choosing the larger food amount were extremely high now (i.e., they received no food at all). Recall that in the original task, the chimpanzees were rewarded at least with the smaller quantity.

Recently, Kralik et al. (2002) tried to replicate the study of Silberberg and Fujita (1996) with cotton-top tamarins (*Saguinus oedipus*). Remarkably, increasing the costs of choosing the larger food quantity did not help tamarins. They constantly picked the larger food quantity instead of the smaller one, even though they received no food at all. Kralik et al. investigated whether the tamarins' failure was caused by the saliency of the food. Results revealed that decreasing the incentive value by closing the experimenter's hands so that the food could not be seen anymore did not change task performance. The tamarins still remembered the location of the larger food quantity. When the tamarins had to choose between two different colors, each associated with a different food amount, they started choosing the smaller food quantity but only if they were not reinforced (i.e., received no rewards) after choosing the color associated with the larger food quantity.

Increasing the costs or manipulating food visibility did not improve test performance for another species of New World monkeys, the squirrel monkey (*Saimiri sciureus*; Anderson et al., 2000). These findings coincide with the results for Japanese macaques described earlier (Silberberg & Fujita, 1996). However, after the authors introduced a modified procedure based on receiving no rewards and a time-out period for selecting the larger quantity, the monkeys successfully performed the "four leads to zero" version of the task as well as the original version without the modified procedure. Performance was maintained when novel array size pairs were presented. When tested after an 8-month delay, the monkeys continued to show inhibitory control by reliably selecting the smaller quantity of the one-versus-four combination within two sessions (Anderson, Awazu, & Fujita, 2004). When novel food array combinations were presented, there was transposition at the group level, though individual differences in bias toward the smaller quantities became apparent (Anderson et al., 2004). Through use of a modified procedure, lemurs have also been shown to be able to choose a smaller food amount when presented with two stimulus arrays of one and four raisins (Genty et al., 2004). The lemurs showed transposition toward the original task without a modified procedure as well as to novel array combinations. When tested after several months, 6 out of 10 lemurs retained a significant preference for the smaller food array.

A recent reversed contingency study by Shumaker et al. (2001) revealed that 2 orangutans were able to optimize performance without any task adaptations. Shumaker et al. presented subjects with pairs of visible quantities of grapes ranging from one to six items. They found that orangutans, unlike chimpanzees, solved the task with relative ease. One of the orangutans reliably picked the smaller food quantity after 7 sessions, and the other did so after 14 sessions. The performance of the orangutans was even above the level achieved by the chimpanzees tested with Arabic numbers in the studies of Boysen and Berntson (1995) and Boysen et al.

(1999). Shumaker et al. hypothesized that the difference between chimpanzees and orangutans might have been produced by the extreme differences in their social systems. In contrast to chimpanzees, orangutans have an extended social system, and because of their solitary existence, they face relatively minor food competition (Goodall, 1986; van Schaik & van Hooff, 1996). According to Shumaker et al., chimpanzees might therefore be more impulsive and, as a consequence, have a harder time optimizing performance on a reversed contingency task.

In summary, to this point several studies have investigated reversed contingency task performance in several nonhuman primate species. These studies have shown that chimpanzees, Japanese macaques, squirrel monkeys, lemurs, and cotton-top tamarins are unable to pass the reversed contingency task (i.e., to select the smaller quantity to net the larger one) unless special manipulations are implemented, such as increasing the cost of selecting the largest quantity by introducing a zero gain contingency or replacing the actual foods with symbols. Recently, however, Shumaker et al. (2001) reported that 2 orangutans solved the task without the use of any modified procedures or symbols. However, the sample size ($N = 2$) of their study makes an explanation at the species level preliminary. As suggested by Shumaker et al., the data can also be explained by individual differences such as differential exposure to cognitive testing.

Therefore, our first goal in this study was to use the same procedure to investigate the performance of chimpanzees and orangutans with comparable experimental backgrounds. On the basis of the study of Shumaker et al. (2001), we expected orangutans to optimize performance and inhibit their preference to choose the larger food quantity, whereas on the basis of the results of Boysen and Berntson (1995), we expected chimpanzees to continue choosing the larger one.

The second goal of this study was to investigate the capabilities of bonobos and gorillas on a reversed contingency task, because obtaining data from all great ape species is essential to make inferences about the evolution of human and ape cognitive skills (see Beck, 1982; Parker, Mitchell, & Miles, 1999). This is particularly important for reversed contingency performance, because knowing more about inhibitory control in great apes may help determine whether the differential problem-solving abilities across the great ape species might be related to differences in inhibitory control rather than interspecific differences in conceptual understanding or other skills.

The third aim of this study was to investigate the joint effect and the relative importance of two variables on reversed contingency performance: food quantity and the incentive value of visible food. Both variables have been shown to affect the performance of chimpanzees and monkeys (e.g., Boysen et al., 1996; Kralik et al., 2002). To investigate the effect of food quantity, we used two food quantities: zero versus four raisins and one versus four raisins. We did not use a modified procedure. This means that subjects always received the opposite quantity to the one that they selected. We predicted that the zero-versus-four combination would be more difficult than the one-versus-four combination. Note that in the zero-versus-four combination the task may be particularly difficult because subjects have to select an empty dish. Set against that, however, the costs are higher when subjects choose the larger food quantity because they then receive nothing. Although zero is perhaps a special type of quantity, we called the comparison of

zero versus four items and one versus four items a test of food quantity.

To investigate the effect of manipulating the incentive value of the food, we used a procedure similar to that used by Kralik et al. (2002) and presented subjects with two types of trials: visible arrays and color-associated covered arrays. In the visible trials, subjects were able to see the actual food quantities, whereas in the color-associated covered trials, subjects were presented dishes covered with colored lids representing the various quantities. Subjects acquired the association between colors and quantities during a training phase prior to the reversed contingency test. We hypothesized that decreasing the incentive value of the arrays by covering them with colored lids that were associated with the food quantities would make the task easier for the apes because it would reduce the prepotent tendency to select the larger amount. We therefore called this a test of food visibility.

Method

Subjects

Four chimpanzees (*Pan troglodytes*; age range: 9–22 years), 5 orangutans (*Pongo pygmaeus*; age range: 5–15 years), 4 bonobos (*Pan paniscus*; age range: 6–20 years), and 5 gorillas (*Gorilla gorilla*; age range: 5–25 years) housed at the Wolfgang Köhler Primate Research Center (WKPRC) in the Leipzig Zoo (Leipzig, Germany) participated in the study (see Table 1). This study originally included 7 orangutans, but 2 of them were dropped because they appeared to be unable to learn the association between colors and food quantities (see *General Performance* section below). Prior to this study, all of the subjects had participated in various experimental studies on cognition. One of these studies had shown that all of the apes at the WKPRC reliably selected the larger of two food quantities (Hanus, Call, & Tomasello, 2003). Subjects were not food or water deprived at any time.

Table 1
Species, Name, Age, Sex, and Rearing History for Each Subject in the Study

Species and name	Age (years)	Sex	Rearing history
Chimpanzee			
Dorien	22	F	Nursery
Jahaga	10	F	Mother
Sandra	10	F	Mother
Frodo	9	M	Mother
Gorilla			
Bebe	24	F	Unknown
Ndiki	25	F	Unknown
Viringika	8	F	Mother
Nkwango	7	M	Mother
Ruby	5	F	Mother
Bonobo			
Joey	20	M	Nursery
Ulindi	9	F	Mother
Limbuko	7	M	Nursery
Kuno	6	M	Nursery
Orangutan			
Dokana	14	F	Mother
Pini	15	F	Mother
Walter	14	M	Mother
Toba	9	F	Mother
Padana	5	F	Mother

Note. F = female; M = male.

During the period of testing, all apes received their complete daily diet consisting of fresh fruits, vegetables, meat, bread, and banana pellets.

Materials

We presented food quantities in open white dishes (5 cm in diameter) or covered colored dishes (colors: orange, purple, and black; dish: 5 cm in diameter; cover: 10 cm in diameter) placed on a platform (90 cm × 40 cm). The platform rested on top of a small table (90 cm × 40 cm × 70 cm) and could slide back (in the direction of the experimenter) and forth (in the direction of the subject). The table was flush against the Plexiglas panel of the subject's cage. The panel had three holes through which the subject could stick its finger to make a choice. We used an opaque barrier (90 cm × 40 cm) to prevent subjects from observing the baiting process. As rewards, the apes received raisins or quarters of grapes depending on the subjects' preferences. Raisins and grapes are highly favored food items by the apes at the WKPRC.

Design

The design was a split plot design with food visibility (visible or covered), food quantity (0–4 or 1–4), and block (Block 1: Sessions 1–10, Block 2: Sessions 11–20, or Block 3: Sessions 21–30) as the within-subject factors and species (chimpanzee, orangutan, bonobo, or gorilla) as the between-subjects factor. There were two color training phases in which subjects learned the association between colored dish lids and food quantities. Thereafter, they were presented 30 test sessions each containing four types of trials resulting from all possible Visibility × Quantity combinations. The four types of trials were presented in random order.

Procedure

Each subject was tested individually in an observation room. We presented subjects with two color training phases and a test phase. In the color training phases, subjects learned the association between colored dishes and three corresponding food quantities (zero, one, and four raisins). They were supposed to choose the color associated with the larger food quantity. After the color associations were acquired, subjects advanced to the test phase, in which they were required to choose the smaller of two food quantities in order to receive the larger quantity.

Color training phases. Color training took place in two phases. In the first phase, we presented subjects with three baited colored dishes (zero, one, or four raisins) covered with their lids. The possible combinations between color of the dishes (orange, purple, or black) and food quantity (zero, one, or four raisins) were counterbalanced within species. Each dish was placed 20 cm away from the Plexiglas panel, in front of one of the three equidistant holes in the panel. We counterbalanced the location of the dishes (left, middle, or right) across trials by using a randomization program. The experimenter baited the dishes behind the opaque barrier, removed the barrier, looked at the subject, and pushed the platform against the Plexiglas as soon as the subject was looking at the task setup. As soon as the experimenter began to push the platform forward, she shifted her gaze from the subject to the middle hole in the Plexiglas panel. If subjects did not respond after 5 min, the session ended and was completed on another day. After the subject made a choice, the experimenter looked up, opened the chosen dish, and then pulled the other two dishes backward simultaneously while opening them in the same movement. Subjects were rewarded with the content of the chosen dish (zero, one, or four raisins) by introducing the raisins one by one through the Plexiglas hole in front of the chosen dish. Each session contained 12 trials (2 trials for each of the six location orders). Subjects moved to the next color training phase when they chose the color associated with the largest food quantity (i.e., four raisins) on 12 trials (100%) of one session. Otherwise, we administered additional sessions until the criterion was met.

In the second phase, we presented subjects with pairs of dishes and used the same basic presentation procedure as in the previous phase with the exception that both dishes were now presented in the middle of the platform. Before the experimenter pushed the platform against the Plexiglas, the dishes were pulled apart until they reached the position of the left and right Plexiglas holes. Subjects were rewarded by introducing the chosen quantity through the center hole in the Plexiglas panel. The order in which pairs were presented and the location (left or right) at which dishes were placed were counterbalanced across trials. Each of the three possible pairings (1–0, 1–4, or 0–4) occurred equally often. Each session contained 12 trials (2 trials for each of the six color orders). If subjects chose the dish containing the larger food quantity in 100% of the trials on three consecutive sessions, they advanced to the test phase.

Test phase. Food quantities were presented in pairs according to the procedure of the second training phase. Subjects were presented with the following four conditions (see Figure 1): (a) visible 0–4: two open white dishes with zero and four raisins; (b) visible 1–4: two open white dishes with one and four raisins; (c) covered 0–4: two covered dishes with colored lids, one associated with zero raisins and the other with four raisins; and (d) covered 1–4: two covered dishes with colored lids, one associated with one raisin and the other with four raisins.

The order of presentation of each type of trial and the locations at which the dishes were placed were counterbalanced across trials with the constraint that the larger food quantity was never in the same location more than three times in a row. After the subject made its choice, the experimenter pulled back the chosen dish, pushed the nonchosen dish forward simultaneously, and opened both dishes at the same time in the case of the covered trials. The subject received the content (or in the case of the colored dishes the associated content) of the dish it did not choose (zero, one, or four raisins). Sessions contained 16 trials (4 trials per condition [4]). Testing stopped after 30 sessions.

When subjects made a double choice (i.e., they simultaneously pointed at both dishes) during color training or test trials, the experimenter pulled the platform back and encouraged the subject to choose a single dish by saying, "You can only choose one." Thereafter, the platform was pushed against the Plexiglas again. If participants did not respond after 5 min, the session ended and was completed on another day. We administered each color training or test session on a different day. The median number of days between sessions was 1. We videotaped all color training and test sessions.
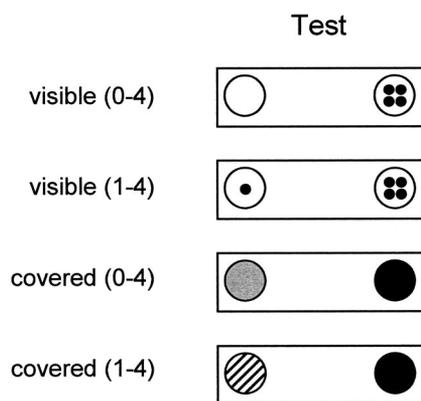
Test

Figure 1. Example of the four conditions used during the test phase. Visible 1–4: two open white dishes baited respectively with one versus four raisins. Visible 0–4: two open white dishes baited with zero versus four raisins. Covered 1–4: two colored dishes associated with one raisin (the purple dish in this example) versus four raisins (the orange dish in this example). Covered 0–4: two colored dishes, one associated with zero raisins (the black dish in this example) and the other with four raisins (the orange dish in this example).

Data Analyses

Subjects' choices during testing were coded twice, online and on tape following the sessions. A choice was counted as correct when the subject chose the smaller food amount. A second coder coded 20% of all taped sessions per species. The intercoder reliability measured by Cohen's kappa was .98 ($p < .01$).

The percentage of correct choices was subjected to a mixed analysis of variance with food visibility (visible or covered), food quantity (0–4 or 1–4), and block (Block 1: Sessions 1–10, Block 2: Sessions 11–20, or Block 3: Sessions 21–30) as the within-subject factors and species (chimpanzee, orangutan, bonobo, or gorilla) as the between-subjects factor. Post hoc analyses were corrected with the Bonferroni–Holm procedure (Holm, 1979).

We also analyzed the speed of acquisition of the reversals in the various conditions. Given the subjects' initial strong predisposition for choosing the larger of the two quantities, we considered the selection of the smaller quantity on at least one out of four trials in 3 consecutive sessions as a valid indication of change. Although one might argue that selecting one trial out of four is not a large change (even if it happens in 3 consecutive sessions), we wanted to capture the first indication that subjects were becoming sensitive to the reversed contingency. We subjected the variable speed of acquisition to a mixed analysis of variance with food visibility (visible or covered) and food quantity (0–4 or 1–4) as factors.

Finally, we analyzed the data individually to investigate two aspects. First, we used one-sample *t* tests to investigate whether subjects selected the correct alternative significantly above chance level after pooling together the last 10 testing sessions. We considered that a subject passed or failed a given condition depending on whether it was significantly above or below chance, respectively. Additionally, the results of those subjects that did not exceed chance levels (either above or below) were deemed inconclusive. Because of the large number of tests involved in this analysis, we set the alpha value at $p < .01$.

Second, we calculated the correlation between the percentage of correct trials and the session number (considering only the last 10 sessions) for each subject (and condition) to investigate whether there was any evidence that subjects were still improving when testing ended. We used the Spearman correlation because of the small number of trials per condition within a session. To guard against reporting spurious correlations, we set the alpha value at $p < .01$.

Results

General Performance

All apes, except for 2 orangutans that were dropped from the study after they failed to master the first color training phase after 29 and 25 sessions, learned the association between the colors and food quantities and chose the larger food quantity on 100% of 3 consecutive sessions in Color Training Phase 2 (see Table 2). Figure 2 depicts the overall percentage of correct trials for each of the four types of trials for all species. Focusing first on the main effects, we found that all species performed at the same level, $F(3, 14) = 0.34$, $p = .80$. Analysis of the within-group main effects revealed that subjects performed better in the 1–4 trials than in the 0–4 trials, $F(1, 14) = 74.57$, $p < .01$, and in the covered trials than in the visible trials, $F(1, 14) = 15.15$, $p = .002$. Likewise, subjects significantly improved their performance across the session blocks, $F(2, 28) = 42.92$, $p < .01$. Overall, subjects were below chance (50%) in the first session for all conditions, one-sample $t(17) > 4.91$, $p < .01$, in all cases. In contrast, they were above chance in the last session for both 1–4 conditions, visible: $t(17) = 3.22$, $p = .005$; covered: $t(17) = 3.57$, $p = .002$, but not for the

Table 2
Mean Number of Sessions (and Standard Errors) Needed to Learn the Color–Quantity Association in Each Color Training Phase for Each Species
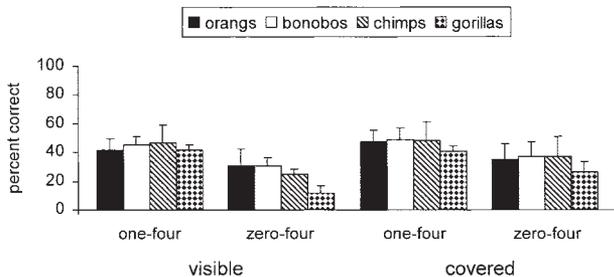
Species	Phase 1		Phase 2	
	M	SE	M	SE
Bonobo	4	1	5	2
Chimpanzee	8	3	7	1
Orangutan	15	5	9	3
Gorilla	18	5	23	12

0–4 conditions, visible: $t(17) = 0.79$, $p = .44$; covered: $t(17) = 2.00$, $p = .062$.
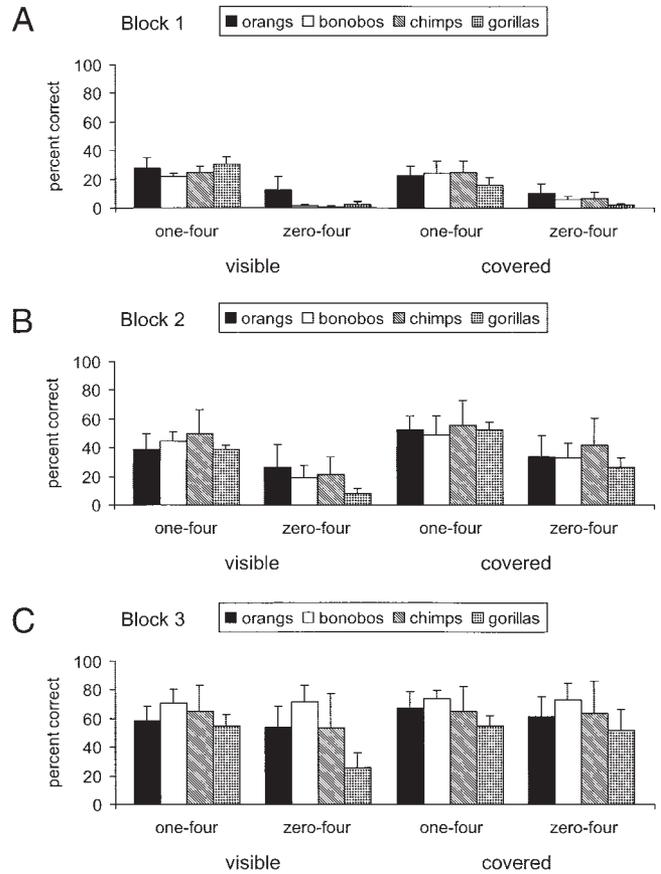
There were several significant interactions between factors as well. There was a Visibility $\times$ Quantity interaction, $F(1, 14) = 5.54$, $p = .034$, suggesting that the effect of visibility was different for the two food quantity combinations. There was also a Quantity $\times$ Block effect, $F(1.73, 28) = 15.09$, $p < .01$, indicating that subjects improved differently for the 0–4 and 1–4 food quantities. Finally, there was a Visibility $\times$ Block effect, $F(2, 28) = 7.70$, $p = .002$, suggesting that subjects improved at different speeds depending on the perceptual features (visible or covered) of the food quantities. Because of these multiple interactions, we separately reanalyzed the data for each block of sessions.

Figure 3A depicts the percentage of correct trials in the first block of 10 sessions for each type of trial for all species. Overall, subjects were below chance in all conditions, $t(17) > 8.77$, $p < .01$, in all cases. There was only a significant effect of quantity, $F(1, 14) = 88.23$, $p < .01$, indicating that subjects performed better in the 1–4 pairs than in the 0–4 pairs (see Table 3).

Figure 3B depicts the percentage of correct trials in the second block of 10 sessions for each type of trial for all species. Overall, subjects were below chance in both 0–4 conditions, visible: $t(17) = 5.89$, $p < .01$; covered: $t(17) = 2.79$, $p = .013$, and they did not differ from chance in the 1–4 conditions, visible: $t(17) = 1.62$, $p = .125$; covered: $t(17) = 0.47$, $p = .64$. There was a significant effect of quantity, $F(1, 14) = 63.06$, $p < .01$, and visibility, $F(1, 14) = 20.16$, $p = .001$. This result indicates that subjects performed better in the 1–4 pairs than in the 0–4 pairs and additionally performed better when presented with covered colored dishes associated with the food quantities (see Table 3).

Figure 3. Mean percentage of correct trials ($+SE$) as a function of condition and species for the (A) first, (B) second, and (C) third block of 10 sessions (each block represents 40 trials per condition). orangs = orangutans.

Figure 3C depicts the percentage of correct trials in the third block of 10 sessions for each type of trial for all species. There was a significant effect of quantity, $F(1, 14) = 7.40$, $p = .017$; visibility, $F(1, 14) = 7.19$, $p = .018$; Quantity $\times$ Visibility, $F(1, 14) = 5.40$, $p = .036$; and Quantity $\times$ Visibility $\times$ Species, $F(3, 14) = 3.63$, $p = .04$ (see Table 3). Because of the three-way interaction, we separately reanalyzed the performance of each species. Orangutans were affected by visibility, $F(1, 4) = 10.12$, $p = .034$; that is, they performed better when presented with covered colored dishes that were associated with food quantity. Gorillas were affected by food quantity, $F(1, 4) = 8.09$, $p = .047$; that is, they performed better in the 1–4 pairs than the 0–4 pairs, whereas chimpanzees and bonobos were unaffected by any of the factors, $F(1, 7) < 1.88$, $p > .21$, for all factors (separately analyzing chimpanzees and bonobos produced comparable results).

Speed of Acquisition

We also analyzed how long subjects took to overcome their strong predisposition to select the larger of two quantities. Figure 4 presents the mean session number in which subjects were correct in at least one trial per session for 3 consecutive sessions for the first time. Subjects initiated their reversals earlier with the 1–4

Figure 2. Mean percentage of correct trials ($+SE$) as a function of condition and species. orangs = orangutans.

Table 3

Analysis of Variance for Species, Food Quantity, Food Visibility, and Their Interactions in Each Block of 10 Sessions

	Block 1		Block 2		Block 3	
Variable	*F*	*p*	*F*	*p*	*F*	*p*
Species	0.35	.79	0.18	.91	0.64	.60
Quantity	88.23**	.00	63.06**	.00	7.40*	.02
Visibility	0.33	.57	20.16**	.00	7.19*	.02
Species × Quantity	0.79	.52	1.12	.38	1.73	.21
Species × Visibility	1.60	.23	0.31	.82	0.70	.57
Quantity × Visibility	2.49	.14	1.29	.28	5.40*	.04
Species × Quantity × Visibility	0.51	.68	0.88	.48	3.63*	.04

$* p < .05.$ $** p < .01.$

pairs than with the 0–4 pairs, $F(1, 14) = 53.33$, $p < .01$. In addition, there was a significant Quantity × Visibility interaction, $F(1, 14) = 7.90$, $p = .014$. All other main effects and interactions were nonsignificant, including those involving species. Post hoc analyses indicated that with the 1–4 combination, subjects initiated their reversals earlier in the visible condition than in the covered condition, $t(17) = 2.68$, $p = .016$, but no significant differences were detected between the visible and covered conditions with the 0–4 combination, $t(17) = 1.53$, $p = .15$. Thus, subjects initiated their reversals first in the visible 1–4 combination, next in the covered 1–4 combination, and finally in those conditions involving the 0–4 combination regardless of the visibility of the food arrays.

Individual Differences

Table 4 presents the percentage of correct trials per subject in the third block of trials for all conditions. There were large individual differences in performance. Two chimpanzees (Sandra and Dorien), 2 bonobos (Ulindi and Limbuko), and 2 orangutans (Walter and Dokana) passed all conditions. That is, they performed above chance in each of the conditions in the third block of 10 sessions. In addition, 1 gorilla (Viringika) passed all conditions except the visible 0–4 condition, and 1 bonobo (Kuno) passed the covered 1–4 condition. In contrast to those successful subjects, 1

chimpanzee (Jahaga) failed all conditions. That is, the subject did not optimize performance (above chance) in the third block of 10 sessions. One orangutan (Pini) and 1 gorilla (Ndiki) each failed two conditions (those including the zero in the pair). One chimpanzee (Frodo) and 2 gorillas (Ruby and Nkwango) failed the visible 0–4 condition. Finally, 2 orangutans (Toba and Padana) and 1 bonobo (Joey) neither passed nor failed, and 1 gorilla (Bebe) failed the visible 0–4 condition but passed the covered 0–4 condition.

Table 4 also presents those cases that displayed positive correlations between the percentage of correct trials and the session number. There were 7 cases with significant positive correlations denoting an improvement over trials and no cases with significant negative correlations. Two of those cases corresponded to an orangutan (Walter) and a gorilla (Bebe) whose performance was already above chance in those conditions. The remaining 5 cases occurred in subjects that were not above chance at the end of testing (1 chimpanzee, 1 gorilla, and 1 orangutan). The case of the orangutan (Padana) was particularly noticeable because she showed evidence of improving in three of the four conditions when testing ended. None of the subjects that were significantly below chance in particular conditions (see Table 4) showed any signs of improvement in any of those conditions, although 1 chimpanzee (Frodo) and 1 gorilla (Bebe) showed improvement in other conditions.

Discussion

In this study, we investigated whether there were differences between the four species of great apes in reversed contingency task performance. Results indicated a comparable performance across species. In general, all great ape species were able to indicate the smaller quantity to get the larger quantity to some extent. However, performance varied depending on the two factors investigated: food quantity and incentive value. First, subjects performed better when presented with the 1–4 food quantity combination then when presented with the 0–4 combination. Second, subjects showed superior performance in covered trials as compared with visible trials, particularly for the 0–4 combination. Quantity had a larger impact on performance than the incentive value of the food (i.e., visibility). In addition, there were large individual differences in reversed contingency task performance—in each species, some apes were very proficient at the end of testing, whereas others were
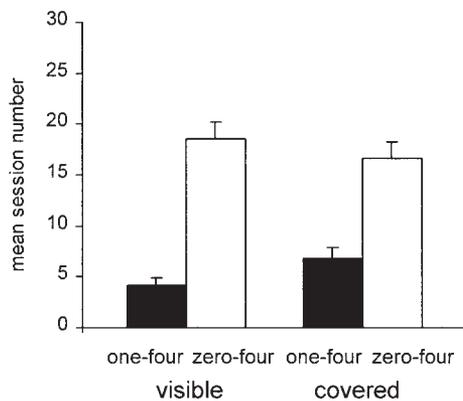

Figure 4. Mean session number (+*SE*) in which subjects were correct in at least one trial for three consecutive sessions for the first time.

Table 4
Mean Percentage of Correct Trials for Each Subject in Each Condition in the Third Block of Trials

Subject's name	Species	Visible		Covered	
		1–4	0–4	1–4	0–4
Sandra	Chimpanzee	100.0**	97.5**	92.5**	97.5**
Dorien	Chimpanzee	90.0**	92.5**	90.0**	97.5**
Walter	Orangutan	90.0**	87.5[a]**	92.5**	95.0**
Limbuko	Bonobo	82.5**	92.5**	85.0**	97.5**
Dokana	Orangutan	72.5**	87.5**	97.5**	95.0**
Ulindi	Bonobo	90.0**	90.0**	77.5**	85.0**
Viringika	Gorilla	85.0**	65.0	82.5**	92.5**
Kuno	Bonobo	62.5	60.0	75.0**	67.5
Bebe	Gorilla	50.0	10.0**	51.0	73.3[a]**
Toba	Orangutan	50.0	45.0	50.0	47.5
Joey	Bonobo	47.5	42.5[a]	57.5	42.5
Frodo	Chimpanzee	45.0[a]	20.0**	60.0	57.5
Padana	Orangutan	45.0	35.0[a]	57.5[a]	42.5[a]
Ruby	Gorilla	50.0	12.5**	57.5	50.0
Nkwango	Gorilla	47.5	30.0**	37.5	37.5
Pini	Orangutan	35.0	15.0**	40.1	25.0**
Ndiki	Gorilla	41.5	10.0**	44.1	5.0**
Jahaga	Chimpanzee	25.0**	2.5**	17.5**	2.5**

[a] Denotes a significant ($p < .01$) positive Spearman correlation between percentage correct and session number.
** $p < .01$ (based on one-sample t test with 50% indicating chance responding).

not. Next, we discuss the implications of each of these findings in turn.

Food Quantity

Subjects performed better in the 1–4 combination than in the 0–4 combination. This effect was already apparent in the first block of trials and persisted throughout testing. Moreover, subjects also reversed faster in the 1–4 combination than in the 0–4 combination. These results are consistent with those of Boysen et al. (1996), who reported that performance was highly related to the relative size disparity between the choice arrays as well as the overall array size (they tested all quantity combinations from zero to six candies). In the study of Boysen et al. (1996), subjects' performance decreased when the disparity between the two food quantities in the presented pair increased, although this effect diminished when the mean array size increased. Those trials involving pairs in which one of the elements was zero were particularly difficult. Similarly, Russell et al. (1991) found that 3-year-old children and children with autism had problems with pointing to nothing in the presence of another salient food amount as well.

However, our results also critically differ from those of Boysen et al. (1996) because some of the individuals in our study, including some chimpanzees, were able to solve the problem even in the visible 1–4 and 0–4 combinations. In contrast, none of the subjects in Boysen et al.'s (1996) study were able to solve this problem (unless they used Arabic numerals). It is interesting to note that several studies have reported that macaques, squirrel monkeys, and lemurs are capable of solving the reversed contingency task involving visible quantities (Anderson et al., 2000; Genty et al., 2004; Silberberg & Fujita, 1996). However, subjects in those studies, unlike those in the current study, invariably required a modified procedure based on changing the reward

contingencies of the task or introducing time-out periods after errors. For instance, Silberberg and Fujita found that macaques were able to point to one to get four raisins but only if subjects received zero raisins (not one raisin, as was the case in Boysen et al.'s [1996] study) when they chose four. This means that Silberberg and Fujita's subjects did not face the problem of pointing to zero to get four (a problem that the data show is much harder than pointing to one to get four). Moreover, subjects were punished with no reward when choosing the four-raisin alternative. Therefore, macaques faced a simpler problem than did the apes in the current study.

There are two studies that obtained positive results without implementing a modified procedure. First, Murray, Kralik, and Wise (2005) found that when rhesus macaques (*Macaca mulatta*) had to choose between one and four pieces of food, they all initially tended to select the larger food quantity. However, subjects were able to learn the standard version of the reversed contingency task when presented with a substantial number of trials (mean number of trials to criterion = 1,087; range = 340–2,700). In comparison with the rhesus macaques, the subjects in the present study improved more quickly and needed fewer trials to optimize or improve their performance.

Second, Shumaker et al. (2001) presented orangutans with the original task introduced by Boysen and Berntson (1995): pairs of visible quantities ranging from one to six items. Unfortunately, Shumaker et al., following Boysen and Berntson (1995), did not use zero items in any of their pairs. Nevertheless, they found that orangutans, unlike the chimpanzees tested by Boysen and Berntson (1995), solved the task with relative ease. One of the orangutans reliably picked the smaller food quantity after 7 sessions, and the other did so after 14 sessions. This result is comparable to the performance of some orangutans included in the current study.

One puzzling aspect of Shumaker et al.'s (2001) results is that, unlike our current and Boysen et al.'s (Boysen & Berntson, 1995; Boysen et al., 1996, 1999, 2001) results, none of their orangutans showed an initial strong preference for the larger quantity—a preference that has been documented not only in reversed contingency studies but also in other studies involving the choice between two nonequal food quantities (e.g., Call, 2000; Dooley & Gill, 1977; Menzel & Draper, 1965; Rumbaugh, Savage-Rumbaugh, & Hegel, 1987; Suda & Call, 2004). This lack of preference for the larger of two quantities means that the orangutans tested by Shumaker et al. did not have to overcome the preference for the larger quantity, which is a key component of the reversed contingency task as it relates to inhibiting a prepotent response tendency of selecting the larger quantity.

Food Visibility

Manipulating the incentive value of the food also had an effect on performance. Subjects showed superior performance when presented with covered colored dishes associated with the different food quantities, particularly for the 0–4 quantity combination. However, this effect was not apparent until the second block of trials. One possible explanation for this result is the extensive color training that preceded our test. Subjects were extensively trained on choosing the color associated with the larger food quantity in two phases. It probably took subjects at least 80 trials (10 sessions × 8 color trials) to be able to overcome this response bias. These results would be consistent with the study of Kralik et al. (2002), who reported that cotton-top tamarins did not benefit from choosing between two colors instead of two visible food quantities after 120 trials. As in the present study, subjects might have needed more trials to overcome the bias toward the color associated with the larger food quantity. Even then, their subjects were able to choose the smaller food quantity in covered trials only after increasing the costs (receiving nothing when choosing the larger quantity).

In contrast to the first block of sessions, subjects performed better in covered trials than in visible trials in the second block of sessions for both food quantity combinations, even though subjects still performed better in the 1–4 combination than in the 0–4 combination (see previous section). Subjects continued to improve their performance in the third block of sessions, and the gap across conditions diminished substantially except for the visible 0–4 combination that still lagged significantly behind the other three conditions. The improvement detected during the second block of sessions for covered trials is consistent with the results of Boysen and Berntson (1995), who found a significant performance improvement when chimpanzees received a conceptual representation of the food quantities rather than the quantities themselves. In the study of Boysen and Berntson (1995), several quantities were represented by Arabic numerals. In the current study, quantities were represented by three colors, which is a much simpler association than Arabic numerals. Similarly, children also perform better on a delay of gratification task when presented with pictures of candies instead of real candies (Mischel, Shoda, & Rodriguez, 1989). However, unlike Boysen and Berntson (1995), this improvement over trials was gradual, not sudden. Indeed, by applying an ABAB design, Boysen and Berntson (1995) found that subjects' performance immediately increased when they were presented with Arabic numerals instead of visible food quantities.

Perceptual Versus Cognitive Processes

In explaining their results regarding the difference between quantity-based and symbol-based trials, Boysen and Berntson (1995) invoked two opposing processes: a perceptual process and a cognitive process. The perceptual process is responsible for the subjects' predisposition to prefer a larger food quantity over a smaller food quantity. The cognitive process is responsible for subjects' learning the rule of choosing the smaller quantity to receive the larger one. According to Boysen and Berntson (1995), the perceptual process might have hindered the expression of the cognitive process. That is, subjects might have learned the rule of choosing the smaller food quantity when presented with real food quantities, but this knowledge could not be expressed because of an interfering predisposition to choose the larger quantity caused by the perceptual process. The fact that performance immediately increased when the interference effect of the perceptual process was eliminated by presenting them with Arabic numerals supports this explanation. Our data suggest that even when the visual access to the quantities is removed in covered trials, subjects still experience an influence of the actual quantities although via the representation—not the direct perception—of those quantities. Therefore, one can hypothesize that visible trials and covered colored trials correspond to three different levels of processing of information of quantities: perceptual, representational, and conceptual.

Reversed Contingency and Reversal Learning

The present reversed contingency trials with covered dishes have much in common with those of a reversal discrimination learning task, and this may be an explanation for the gradual reversal. In reversal discrimination problems, one of two stimuli is rewarded for several trials. Thereafter, the contingency is changed. Subjects are no longer rewarded for selecting the previously rewarded stimulus but need to choose the previously unrewarded stimulus. Several studies have reported that apes perform well on reversal discrimination learning (Davis & Markowitch, 1978; Essock & Susan, 1978; Gill & Rumbaugh, 1974; Rumbaugh, 1971; Rumbaugh & Pate, 1984a, 1984b; Schusterman, 1962, 1964; Tomonaga, 1999). So it is not surprising that some apes could solve the reversed contingency problem with covered dishes in the present study as well. In effect, what this finding may imply is that by the end of testing, successful subjects had come to associate the zero- or one-item color with the gain of four rewards and the four-item color with zero or one.

However, there is one piece of evidence that suggests that the present task is not entirely equivalent to reversal discrimination learning. Although the 0–4 pair is more similar to the traditional reversal discrimination task because subjects get no reward for selecting the incorrect alternative (whereas they get rewarded with one piece in the 1–4 pair), the 0–4 pair was harder than the 1–4 pair. Because the food quantities were not visually available in covered trials, this suggests that subjects were using a representation of the quantities in solving the problem. Otherwise, they would simply respond as they would in a reversal color discrimination task, and there would be no reason to expect differences in

the visibility effect between the 0–4 and 1–4 pairs. The fact that the differences between quantities in the covered condition closely reproduced those found in the visible condition lends further support to the hypothesis that quantity representation played a key role in the covered trials.

Species and Individual Differences

In this study, we found no major interspecific differences in performance. Contrary to the expectations based on a previous study (Shumaker et al., 2001), orangutans performed no better than the other species. Although Shumaker et al. recognized the possibility that their results may have reflected individual differences, based on the consistency of their 2 subjects, they favored the idea that species differences were a more probable explanation for the disparity between their results and those of Boysen and colleagues (Boysen & Berntson, 1995; Boysen, Berntson, Hannan, & Cacioppo, 1996; Boysen, Berntson, & Mukobi, 2001; Boysen, Mukobi, & Berntson, 1999).

Our current results argue against a species-based explanation because we found no systematic interspecies differences, and this finding was paired with striking individual differences. Through the use of identical methods with chimpanzees and orangutans (and gorillas and bonobos), some individuals of each species were able to solve this task including the hardest problem represented by the visible 0–4 condition. Two orangutans performed extremely well, 1 of them already in the first block. Another orangutan consistently picked the larger food quantity in most of the trials, reaching her peak performance in the last three sessions. Two other orangutans showed severe inhibition problems. In contrast to the findings of Boysen et al. (Boysen & Berntson, 1995; Boysen et al., 1996, 1999, 2001), there were 2 chimpanzees (and 2 bonobos and 1 gorilla) that were able to solve the reversed contingency task. Another chimpanzee performed extremely poorly, showing a comparable performance to other chimpanzees from previous studies (Boysen & Berntson, 1995).

Thus, in general, all great ape species performed in a similar manner throughout testing, and large individual differences rather than species differences are likely to underlie the putative interspecific differences postulated in a previous study. The only hint of potential interspecific differences in the current study came in the last block of trials. At that stage, orangutans seemed particularly affected by food visibility, whereas gorillas were affected by food quantity. Chimpanzees and bonobos were not affected by either factor anymore. Yet, note that the species still did not differ in overall performance.

An inspection of the subjects' initial and final performance allowed us to distinguish three kinds of individuals. First, there were those subjects that performed above chance levels. These subjects had overcome their initial tendency to select the larger of two quantities in order to optimize their gain by selecting the smaller of the two quantities. Second, there were those subjects that performed below chance levels. These subjects still preferred to select that larger quantity. Finally, there were those subjects that performed neither above nor below chance levels. It is hard to interpret their performance. One possibility is that although they had initially preferred the larger quantity, they had given up trying to net the larger quantity, and they were choosing randomly. Although this may explain the performance of some subjects, it

does not apply to all of them. In particular, there were several cases in which we detected a significant performance improvement in the last 10 sessions. This means that although these subjects did not reach above chance levels at the end of testing, we cannot say that they failed the task either. Presumably, if these subjects had received additional trials, they may have reached above chance levels. Indeed, Murray et al. (2005) showed that rhesus macaques were able to master the task without a modified procedure when presented with a substantial number of trials (ranging from 400 to 2,800) and suggested that other primate species may show comparable abilities provided they receive enough trials.

The inability of some subjects to optimize performance in this task does not necessarily imply an inherent inability to suppress the prepotent response of reaching for the larger food quantity. Kralik (2005) showed that although cotton-top tamarins failed to solve a reversed contingency task in which reaching for one piece of food was rewarded with three pieces whereas reaching for three produced no food at all, they did solve the problem when reaching for one piece was rewarded with only one. It is not entirely obvious why this reduction in the reward for choosing one should have made the task soluble, but because the choice was still between three pieces and one piece of food, if the tamarins' initial failure reflected an inability to inhibit a prepotent response, they should have continued to fail.

Alternative Explanations and Future Directions

Besides individual differences, there are at least three other alternatives that may explain the discrepancies found in the chimpanzees' performance across laboratories. First, the subjects included in the current study had received only one relative numerousness task prior to the reversed contingency test. In contrast, several of the chimpanzees tested by Boysen and colleagues had extensive experience in numerical tasks, both with visible quantities and Arabic numerals (e.g., Boysen & Berntson, 1989, 1995). It is conceivable that this additional experience may have hindered their reversal performance with visible items and, at the same time, may have helped them with Arabic numerals. Second, the chimpanzees tested by Boysen and colleagues received all possible combinations between zero and six items, whereas we used only two combinations in the current study. Multiple combinations may have been harder to master because there is greater variation between pair quantities. This would make solving the problem by using a specific rule for each single pair more demanding than in our study, in which there were only two combinations. However, recall that Shumaker et al. (2001) also used multiple combinations and still found positive results. Third, in the current study, visible and covered trials were presented in the same session, whereas chimpanzees in previous studies experienced them in alternating sessions. This may have helped our subjects to solve the visible trials by "hitchhiking" on the easier covered trials. This explanation, however, does not fit our current data because subjects solved the visible 1–4 pair faster than any other pair. Moreover, recall that the effect of visibility became apparent only after the first block of trials.

As suggested by Kralik et al. (2002), future studies on inhibitory control, like the reversed contingency task, are important for the understanding of interspecific differences in problem solving. Knowing that chimpanzees (who are not extremely number

trained) perform no different from other great ape species on a reversed contingency task eliminates future inhibition explanations for species differences on problem-solving tasks involving food selection. Reversed contingency task performance further provides researchers with information on differences between individuals (Bell, 1998). It is imperative to take into account these inhibition problems when designing experiments investigating cognitive skills.

In conclusion, in this study we have shown that some individuals of each great ape species, not just orangutans, can solve the reversed contingency task without the help of symbolic devices or a modified procedure based on changing the reward contingencies and adding a time-out period for incorrect responses. This study has also confirmed that both the quantity and the visibility of the pairs have a significant effect on performance. In particular, subjects performed better when (a) the disparity between quantities was smaller (and one rather than zero was used as the smaller quantity) and (b) the quantities were not directly visible. Finally, our results indicate important individual differences but no systematic species differences. These results should be confirmed with larger samples.

References

Anderson, J. R., Awazu, S., & Fujita, K. (2000). Can squirrel monkeys (*Saimiri sciureus*) learn self-control? A study using food array selection tests and reversed reward contingency. *Journal of Experimental Psychology: Animal Behavior Processes, 26,* 87–97.

Anderson, J. R., Awazu, S., & Fujita, K. (2004). Squirrel monkeys (*Saimiri sciureus*) choose smaller food arrays: Long-term retention, choice with nonpreferred food, and transposition. *Journal of Comparative Psychology, 118,* 58–64.

Beck, B. (1982). Chimpocentrism: Bias in cognitive ethology. *Journal of Human Evolution, 11,* 3–17.

Bell, M. A. (1998). Frontal lobe functioning during infancy: Implications for the development of cognition and attention. In J. E. Richards (Ed.), *Cognitive neuroscience of attention* (pp. 287–316). Mahwah, NJ: Erlbaum.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology, 103,* 23–31.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: Perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 82–86.

Boysen, S. T., Berntson, G. G., Hannan, M. B., & Cacioppo, J. T. (1996). Quantity-based interference and symbolic representations in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes, 22,* 76–86.

Boysen, S. T., Berntson, G. G., & Mukobi, K. L. (2001). Size matters: Impact of size and quantity on array choice by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 115,* 106–110.

Boysen, S. T., Mukobi, K. L., & Berntson, G. G. (1999). Overcoming response bias using symbolic representations of number by chimpanzees (*Pan troglodytes*). *Animal Learning and Behavior, 27,* 229–235.

Call, J. (2000). Estimating and operating on discrete quantities in orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology, 114,* 136–147.

Davis, R. R., & Markowitch, H. (1978). Orangutan performance on a light–dark reversal discrimination in the zoo. *Primates, 19,* 755–759.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the human brain.* New York: Norton.

Diamond, A. (1988). Differences between adult and infant cognition: Is the

crucial variable presence or absence of language? In L. Weiskrantz (Ed.), *Thought without language* (pp. 337–370). Oxford, England: Clarendon Press.

Diamond, A. (1990). Developmental time course in human infants and infant monkeys, and the neural basis of the inhibitory control of reaching. In A. Diamond (Ed.), *The development and neural bases of higher cognitive functions* (pp. 637–676). New York: New York Academy of Sciences.

Dooley, G. B., & Gill, T. V. (1977). Acquisition and use of mathematical skills by a linguistic chimpanzee. In D. M. Rumbaugh (Ed.), *Language learning by a chimpanzee: The Lana project* (pp. 247–260). New York: Academic Press.

Essock, V., & Susan, M. (1978). Comparison of ape and monkey modes of problem solution. *Journal of Comparative and Physiological Psychology, 92,* 942–957.

Genty, E., Palmier, C., & Roeder, J. J. (2004). Learning to suppress responses to the larger of two rewards in two species of lemurs, *Eulemur fulvus* and *E. macaco. Animal Behaviour, 67,* 925–932.

Gill, T. V., & Rumbaugh, D. M. (1974). Learning processes of bright and dull apes. *American Journal of Mental Deficiency, 78,* 683–687.

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior.* Cambridge, MA: Belknap Press.

Hanus, D., Call, J., & Tomasello, M. (2003). Quantity-based judgments by orangutans (*Pongo pygmaeus*), gorillas (*Gorilla gorilla*) and bonobos (*Pan paniscus*). *Folia Primatologica, 74,* 196–197.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65–70.

Kralik, J. D. (2005). Inhibitory control and response selection in problem solving: How cotton-top tamarins (*Saguinus oedipus*) overcome a bias for selecting the larger quantity of food. *Journal of Comparative Psychology, 119,* 78–89.

Kralik, J. D., Hauser, M. D., & Zimlicki, R. (2002). The relationship between problem solving and inhibitory control: Cotton-top tamarin (*Saguinus oedipus*) performance on a reversed contingency task. *Journal of Comparative Psychology, 116,* 39–50.

Menzel, E. W., & Draper, W. A. (1965). Primate selection of food by size: Visible versus invisible rewards. *Journal of Comparative and Physiological Psychology, 59,* 231–239.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989, May 26). Delay of gratification in children. *Science, 244,* 933–938.

Murray, E. A., Kralik, J. D., & Wise, S. P. (2005). Learning to inhibit prepotent responses: Successful performance by rhesus macaques, *Macaca mulatta,* on the reversed-contingency task. *Animal Behaviour, 69,* 991–998.

Parker, S. T., Mitchell, R. W., & Miles, H. L. (1999). *The mentalities of gorillas and orangutans: Comparative perspectives.* Cambridge, England: Cambridge University Press.

Rumbaugh, D. M. (1971). Evidence of qualitative differences in learning processes among primates. *Journal of Comparative and Physiological Psychology, 76,* 250–255.

Rumbaugh, D. M., & Pate J. L. (1984a). The evolution of cognition in primates: A comparative perspective. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 569–587). Hillsdale, NJ: Erlbaum.

Rumbaugh, D. M., & Pate, J. L. (1984b). Primates' learning by levels. In G. Greenberg & E. Tobach (Eds.), *Behavioral evolution and integrative levels* (pp. 221–240). Hillsdale, NJ: Erlbaum.

Rumbaugh, D. M., Savage-Rumbaugh, S., & Hegel, M. T. (1987). Summation in the chimpanzee (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes, 13,* 107–115.

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The "windows task" as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9,* 331–349.

Santos, L. R., Ericson, B. N., & Hauser, M. D. (1999). Constraints on

problem solving and inhibition: Object retrieval in cotton-top tamarins (*Saguinus oedipus oedipus*). *Journal of Comparative Psychology, 113,* 186–193.

Schusterman, R. J. (1962, August 10). Transfer effects of successive discrimination reversal training in chimpanzees. *Science, 137,* 422–423.

Schusterman, R. J. (1964). Successive discrimination–reversal training and multiple discrimination training in one-trial learning by chimpanzees. *Journal of Comparative and Physiological Psychology, 58,* 153–156.

Shumaker, R. W., Palkovich, A. M., Beck, B. B., Guagnano, G. A., & Morowitz, H. (2001). Spontaneous use of magnitude discrimination and ordination by the orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology, 115,* 385–391.

Silberberg, A., & Fujita, K. (1996). Pointing at smaller food amounts in an analogue of Boysen and Berntson's (1995) procedure. *Journal of the Experimental Analysis of Behavior, 66,* 143–147.

Suda, C., & Call, J. (2004). Piagetian liquid conservation in the great apes (*Pan paniscus, Pan troglodytes,* and *Pongo pygmaeus*). *Journal of Comparative Psychology, 118,* 265–279.

Tomonaga, M. (1999). Establishing functional classes in a chimpanzee (*Pan troglodytes*) with two-item sequential responding procedure. *Journal of the Experimental Analysis of Behavior, 72,* 57–79.

van Schaik, C. P., & van Hooff, J. A. R. A. M. (1996). Toward an understanding of the orangutan's social system. In W. C. McGrew, L. F. Marchant, & T. Nishida (Eds.), *Great ape societies* (pp. 3–15). Cambridge, England: Cambridge University Press.