

How the Great Apes (*Pan troglodytes*, *Pongo pygmaeus*, *Pan paniscus*, *Gorilla gorilla*) Perform on the Reversed Reward Contingency Task II: Transfer to New Quantities, Long-Term Retention, and the Impact of Quantity Ratios

Jana Uher

Max Planck Institute for Evolutionary Anthropology and
Humboldt-University Berlin

Josep Call

Max Planck Institute for Evolutionary Anthropology

We tested 6 chimpanzees (*Pan troglodytes*), 3 orangutans (*Pongo pygmaeus*), 4 bonobos (*Pan paniscus*), and 2 gorillas (*Gorilla gorilla*) in the reversed reward contingency task. Individuals were presented with pairs of quantities ranging between 0 and 6 food items. Prior to testing, some experienced apes had solved this task using 2 quantities while others were totally naïve. Experienced apes transferred their ability to multiple-novel pairs after 6 to 19 months had elapsed since their initial testing. Two out of 6 naïve apes (1 chimpanzee, 1 bonobo) solved the task—a proportion comparable to that of a previous study using 2 pairs of quantities. Their acquisition speed was also comparable to the successful subjects from that study. The ratio between quantities explained a large portion of the variance but affected naïve and experienced individuals differently. For smaller ratios, naïve individuals were well below 50% correct and experienced ones were well above 50%, yet both groups tended to converge toward 50% for larger ratios. Thus, some apes require no procedural modifications to overcome their strong bias for selecting the larger of 2 quantities.

Keywords: inhibition, perseveration, quantity discrimination, relative numerosness, reversal learning

Inhibitory control processes play a central role in problem solving behavior. When direct approach to a food source entails competitive or even predation risks, it can be advantageous to suppress impulses and to pursue alternatives. In the presence of a dominant group member, chimpanzees (*Pan troglodytes*) can inhibit reaching directly for hidden food (Menzel, 1974). Similarly, when mating competition is high, Rhesus monkeys (*Macaca mulatta*) can suppress their copulation calls (Hauser, 1993); and vervet monkeys (*Cercopithecus aethiops*) can suppress their alarm calls in the presence of nonkin as opposed to kin (Cheney & Seyfarth, 1985). Thus, suppressing certain prepotent responses can be advantageous and adaptive. A failure to control such impulses can also prevent subjects from succeeding in problem solving tasks (Deacon, 1997; Diamond, 1990).

A standard paradigm to test inhibitory skills is the reversed reward contingency task. In this task, subjects are rewarded with the nonchosen alternative when presented with two different quantities. Thus, to maximize their rewards subjects have to inhibit their prepotent response to reach for the larger quantity and select the smaller one instead. Different species have been studied in the

reversed reward contingency task including sea lions (*Zalophus californianus*, Genty & Roeder, 2006), lemurs (*Eulemur fulvus*, *E. macaco*, Genty, Palmier, & Roeder, 2004), several monkey species (cotton-top tamarins, *Saguinus oedipus*, Kralik, Hauser, & Zimlicki, 2002; squirrel monkeys, *Saimiri sciureus*, Anderson, Awazu, & Fujita, 2000; Japanese macaques, *Macaca fuscata*, Silberberg & Fujita, 1996; Rhesus macaques, *Macaca mulatta*, Murray, Kralik, & Wise, 2005; mangabeys, *Cercocebus torquatus lunulatus*, Albiach-Serrano, Guillén-Salazar, & Call, 2007), and all the great apes (chimpanzees, *Pan troglodytes*, orangutans, *Pongo pygmaeus*, bonobos, *Pan paniscus*, gorillas, *Gorilla gorilla*; Vlamings, Uher, & Call, 2006; only chimpanzees, Boysen & Berntson, 1995; Boysen, Berntson, Hannan, & Cacioppo, 1996; Boysen, Berntson, & Mukobi, 2001; Boysen, Mukobi, & Berntson, 1999; and only orangutans, Shumaker, Palkovich, Beck, Guagnano, & Morowitz, 2001).

Although rewarding the nonchosen of two unequal alternatives is the key feature of all reversed reward contingency tasks, their specific demands and procedures vary significantly across studies. There is a growing body of research showing that species that initially showed negative results, such as chimpanzees (Boysen & Berntson, 1995; Boysen et al., 1996) or cotton-top tamarins (Kralik et al., 2002) can master the task successfully with procedural changes or with extensive experience. Thus, the present evidence suggests that focusing on the factors influencing performance rather than on species differences may be more informative.

Some factors influencing performance include food quantity, disparity ratio, and item size of the presented arrays. The quantities used in this task range from combinations between zero to six

Jana Uher, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, and Institute for Psychology, Humboldt-University Berlin; Josep Call, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.

Correspondence concerning this article should be addressed to Jana Uher, Humboldt-Universität zu Berlin, Psychologisches Institut, Rudower Chaussee 18, D-12489 Berlin, Germany. E-mail: uher@primate-personality.net

(Boysen et al., 1996), one to five (Boysen et al., 2001), and one to six items (Boysen & Berntson, 1995; Shumaker et al., 2001) to a selection of only a subset of combinations, typically, one against four and zero against four (Anderson et al., 2000; Genty et al., 2004; Kralik et al., 2002; Murray et al., 2005; Silberberg & Fujita, 1996; Vlamings et al., 2006). Results are inconclusive on whether a larger number of arrays makes the task harder or not. Disparity ratio, by contrast, was shown to cause interference effects on performances. In chimpanzees, it accounted for approximately 41% of variance in performance that deteriorated progressively with higher disparity ratios. Moreover, item size was shown to exert an even more powerful inference effect than disparity ratio causing more severe inhibition problems in chimpanzees (Boysen et al., 1999, 2001).

Various procedures imposing different costs to the selector have been administered to different species. Giving the quantity chosen by the subject to another individual while the subject received the nonchosen quantity did not help chimpanzees to master the task (Boysen & Berntson, 1995). However, increasing the costs of choosing the larger quantity by rewarding the subjects with no food helped Japanese macaques to improve their performances (Silberberg & Fujita, 1996). In contrast, neither squirrel monkeys nor cotton-top tamarins improved their performance with this procedure (Anderson et al., 2000; Kralik et al., 2002). Only after instituting timeout and correction procedures, in which the subjects received the quantities in the same positions on every trial until they chose the smaller quantity, squirrel monkeys and lemurs successfully performed the task (Anderson et al., 2000; Genty et al., 2004; Kralik, 2005). Remarkably, great apes started reversing their choices in one against four combinations first, and finally in those conditions involving zero against four food items. Thus, the increased costs in the latter conditions did not make it easier for great apes to overcome their bias (Vlamings et al., 2006).

The impact of the salience of the food was considered a further factor influencing performances. When food arrays were replaced with Arabic numerals, performance immediately increased significantly in chimpanzees. This suggests that the incentive properties of the actual food array made it hard for subjects to inhibit choosing the larger array, and that conversely the symbolic representation of the different quantities helped increasing performance by processing abstract stimuli (Boysen & Berntson, 1995; Boysen et al., 1999). For cotton-top tamarins, decreasing the incentive value of the food by closing the experimenter's hands so that the food was no longer visible did not alter their performance. However, when they had to choose between two colors that had been associated with different amounts of food, instead of seeing the quantities shortly before covered, they started choosing the smaller food quantity, but only if costs were kept high with no reward for choosing the color associated with the larger food quantity (Kralik et al., 2002). Comparing performances in visible and color associated trials, Vlamings et al. (2006) found that manipulating the incentive value of the food had a significant effect on performance in four great ape species. Subjects showed superior performance in color associated trials particularly for the zero against four combinations even though the impact of food visibility was smaller than the impact of quantity.

Transposition effects to novel quantity combinations became apparent at the group level in squirrel monkeys with individual differences in the bias toward the smaller quantities even after an

8-month delay (Anderson, Awazu, & Fujita, 2004). Lemurs also showed transposition toward the original task without any correction procedure as well as to novel array combinations when being retested after 3 months and even after 18 months, although with poorer performances after 18 months (Genty et al., 2004). Mangabeys were also able to perform successfully after 7 to 10 months had elapsed since the end of the last testing (Albiach-Serrano et al., 2007). It is remarkable that those species retained a significant preference for the smaller food array in retests that took place several months after the initial training. In contrast, chimpanzees appeared not to show any evidence of a transposition effect when Arabic numerals were replaced by food arrays. Boysen et al. (1999) concluded, "clearly, the task interference effect with . . . [the nonsubstitute condition] is an enduring phenomenon that does not appear to yield to extended learning experience or success with . . . [the substitute condition]) in the same task with the same reinforcement contingencies" (p. 233).

More recent, several studies have found positive results without any procedural changes but with a sufficient number of trials. Successful subjects mastered the task after dozens or more common hundreds of trials. Successful sea lions have shown the fastest speed of acquisition taking an average of 123 trials to solve the task (Genty & Roeder, 2006) whereas successful mangabeys and Rhesus macaques took on average 485 and 1,087 trials, respectively (Albiach-Serrano et al., 2007; Murray et al., 2005). Vlamings et al. (2006) reported that great apes solved the task after an average of 323 trials but regular trials were presented concurrently in the same session with trials with colored boxes and this may have facilitated task acquisition. Shumaker et al. (2001) reported a faster acquisition in two orangutans (210 trials on average) but these subjects showed no initial preference for the larger of the two quantities. Unlike sea lions (Genty & Roeder, 2006), some mangabeys and great apes have solved pairs that included zero as one of the quantities (Albiach-Serrano et al., 2007; Vlamings et al., 2006). Several studies have documented that after the initial testing with one or two pairs of quantities some sea lions, mangabeys, squirrel monkeys, and lemurs were able to generalize to novel quantity combinations (Albiach-Serrano et al., 2007; Anderson et al., 2000; Genty et al., 2004; Genty & Roeder, 2006).

In sum, currently there is a growing body of data showing that several primate species and sea lions can solve the reversed reward contingency task even without using any procedural modifications to the original task and can transfer this ability to novel pairs of quantities after a long period of time. However, most of the recent studies have used only a small subset of pairs of quantities, typically one against four, and zero against four (Anderson et al., 2000; Genty et al., 2004; Kralik et al., 2002; Murray et al., 2005; Silberberg & Fujita, 1996; Vlamings et al., 2006). This contrasts with the studies with chimpanzees and orangutans that used multiple pairs of quantities ranging from combinations between zero to six items (Boysen & Berntson, 1995; Boysen et al., 1996, 2001; Shumaker et al., 2001). The use of different numbers of quantities may compromise the comparability across species because using multiple pairs may affect the speed of acquisition, transfer, and retention of the reversed reward contingency task. In particular, mastering and remembering how to respond to multiple pairs (as opposed to just a couple of them) may require more time and effort, partly due to potential interference between the quantities in different arrays. For instance, the quantity four is correct in the six

versus four pair but it is incorrect in the four versus two pair. Moreover, storing information about multiple pairs for a long-retention interval creates a greater memory demand in terms of both storage capacity and potential interference between the stored information.

It is thus conceivable that the different number of quantities used in each study could explain the discrepancies between studies on chimpanzees. Whereas the chimpanzees studied by Boysen and Berntson (1995) had to master multiple pairs, those studied by Vlamings et al. (2006) only had to master two pairs of quantities. Although it is true that the two orangutans tested by Shumaker et al. (2001) solved the task with multiple pairs, they did not show a preference for the larger of the two pairs during their initial training. Therefore, unlike most subjects in other studies, these orangutans did not have to overcome the prepotent response of selecting the larger quantity. With regard to transfer, it is unknown whether great apes tested on a couple of pairs will transfer their skill to multiple-novel pairs and whether they will be able to do so after a prolonged period without engaging in the reversed reward contingency task. Finally, the study of only two pairs does not allow us to explore the effect of ratio between quantities, which has been shown to determine performance in quantity discrimination (Beran, 2001; Call, 2000; Hanus & Call, 2007) and reversed reward contingency performances (Boysen et al., 1995, 1999, 2001).

In this study we attempted to overcome three shortcomings in the literature by investigating the ability of two groups of apes to solve the reversed reward contingency problem with multiple pairs ranging from zero to six items. First, to assess the impact of the number of quantity pairs on the initial acquisition, we compared the performance of a group of naïve apes presented with multiple pairs with the performance of the apes that had been tested with two pairs (0:4 and 1:4) in the Vlamings et al. (2006) study. Furthermore, the apes in the current study received no concurrent

trials in which colored boxes substituted the actual quantities. Second, to investigate a potential transfer from two quantity combinations to multiple-quantity combinations we presented multiple pairs to the apes who had successfully reversed their choices in the Vlamings et al. study. Because these subjects also received the pairs that they had mastered initially a few months earlier, this also allowed us to assess their long-term retention of the task. Third, the inclusion of multiple pairs allowed us to assess the effect of ratio on the acquisition and transfer to new pairs.

Method

Subjects

Six chimpanzees (*Pan troglodytes*) (age range 7 to 28 years), three orangutans (*Pongo pygmaeus*) (age range 7 to 31 years), four bonobos (*Pan paniscus*) (age range 7 to 11 years), and two gorillas (*gorilla gorilla*) (age range 7 to 9 years) housed at the Wolfgang Köhler Primate Research Center (WKPRC) in the Leipzig Zoo, Germany, participated in the present study. This sample composed of two different subsamples that are labeled as experienced and naïve subjects throughout the text (see Table 1 and Table 2). The experienced sample composed of all apes who had performed above chance level in the last study 6 to 19 months ago (Vlamings et al., 2006; except for one gorilla [G–Nk] who did not reach chance levels, but who was the second best in the nonsubstitute condition). The naïve sample composed of apes who did not participate in the last study, and who therefore had no experiences with the reversed reward contingency task.

All subjects had participated in various experimental studies on cognition prior to this and the previous study, of which one had shown that all apes at the WKPRC reliably selected the larger of two food quantities (Hanus & Call, 2007). The subjects were always treated in accordance with ethical principles of noninvasive

Table 1
Species, Name, Experimental Condition, and Months Elapsed Since the Original Test, Sex, Age, and Rearing History of the Subjects

| Species and subject | Name | Experimental condition | Months elapsed | Sex | Age ^a | Rearing history |
|---------------------|-----------|------------------------|----------------|-----|------------------|-----------------|
| Bonobo(B) | | | | | | |
| B-Ku | Kuno | Experienced | 14 | M | 8 | Nursery |
| B-Li | Limbuko | Experienced | 14 | M | 9 | Nursery |
| B-Ul | Ulindi | Experienced | 12 | F | 11 | Mother |
| B-Ya | Yasa | Naïve | — | F | 7 | Mother |
| Chimpanzee (C) | | | | | | |
| C-Do | Dorien | Experienced | 19 | F | 24 | Nursery |
| C-Fi | Fifi | Naïve | — | F | 11 | Mother |
| C-Fk | Fraukje | Naïve | — | F | 28 | Nursery |
| C-Ge | Gertruida | Naïve | — | F | 11 | Mother |
| C-Pa | Patrick | Naïve | — | M | 7 | Mother |
| C-Sa | Sandra | Experienced | 19 | F | 11 | Mother |
| Gorilla (G) | | | | | | |
| G-Nk | N'Kwango | Experienced | 7 | M | 7 | Mother |
| G-Vi | Viringika | Experienced | 6 | F | 9 | Mother/peer |
| Orangutan (O) | | | | | | |
| O-Dk | Dokana | Experienced | 11 | F | 16 | Mother |
| O-Du | Dunja | Naïve | — | F | 31 | Nursery |
| O-Wa | Walter | Experienced | 11 | M | 7 | Mother/peer |

Note. The experienced subjects had solved the task in a previous study (Vlamings et al., 2006). Differences in the time that had elapsed since this study was conducted were due to organizational reasons only. M = male; F = female.

^aGiven in years.

Table 2
Subjects, Experimental Condition, and Acquisition Rates Shown in Performances Across Four Blocks of Each Five Sessions

| Subjects | Mean % of correct trials | | | | No. of significantly correct sessions ^a | | | |
|-------------|--------------------------|---------|----------|----------|--|---------|----------|----------|
| | 1 to 5 | 6 to 10 | 11 to 15 | 16 to 20 | 1 to 5 | 6 to 10 | 11 to 15 | 16 to 20 |
| Experienced | | | | | | | | |
| C-Sa | 85.7 | 94.3 | 97.1 | 98.1 | 4 | 5 | 5 | 5 |
| C-Do | 51.4 | 90.5 | 93.3 | 94.3 | 0 | 5 | 5 | 5 |
| B-Ul | 70.5 | 89.5 | 92.4 | 90.5 | 3 | 5 | 5 | 5 |
| B-Li | 76.2 | 86.7 | 94.3 | 89.5 | 4 | 5 | 5 | 4 |
| B-Ku | 60.0 | 93.3 | 88.6 | 90.5 | 2 | 5 | 4 | 5 |
| O-Wa | 68.6 | 73.3 | 85.7 | 88.6 | 1 | 3 | 4 | 5 |
| G-Vi | 68.6 | 86.7 | 84.8 | 85.7 | 2 | 4 | 5 | 5 |
| O-Dk | 42.9 | 53.3 | 49.5 | 50.5 | 0 | 0 | 0 | 0 |
| G-Nk | 32.4 | 41.9 | 46.7 | 45.7 | 0 | 0 | 0 | 0 |
| Naïve | | | | | | | | |
| C-Ge | 39.0 | 53.5 | 83.8 | 95.2 | 0 | 0 | 4 | 5 |
| B-Ya | 31.4 | 41.9 | 60.0 | 91.4 | 0 | 0 | 1 | 5 |
| C-Fk | 34.3 | 47.6 | 51.4 | 49.5 | 0 | 0 | 0 | 0 |
| C-Fi | 20.0 | 49.5 | 48.6 | 48.6 | 0 | 0 | 0 | 0 |
| O-Du | 16.2 | 13.3 | 36.2 | 46.7 | 0 | 0 | 0 | 0 |
| C-Pa | 14.3 | 13.3 | 27.6 | 24.8 | 0 | 0 | 0 | 0 |

Note. The experienced subjects had solved the task in a previous study (Vlamings et al., 2006). See Table 1 for explanation of names of subjects.

^aAnalyzed using binomial tests.

research. During the period of testing all apes received their complete daily diet consisting of various fresh fruits, vegetables, leaves, cereals, eggs, and meat, and they were never deprived of food or water at any time.

Materials

We presented food quantities in two open white dishes (12 cm/4.7 in. in diameter) placed on a table (90 × 40 cm/35.4 × 15.7 in.), which was flush with the Plexiglas panel of the subjects' cage. The panel had two holes at its bottom (50 cm/19.7 in. apart from another) through which the subject could stick a finger to make a choice. We used an opaque barrier (85 × 35 cm/33.5 × 13.8 in.) to prevent subjects from observing the baiting process. As rewards the apes received monkey pellets or quarters of medium-sized dark grapes depending on the subjects' preferences. Both are highly favored food items for the apes at the research center. The prepared baits were stored in a supply cup held by the experimenter underneath the table throughout the session.

Design

Compared to the last study (Vlamings et al., 2006) we enlarged the number of arrays to 21 and presented all possible combinations of food quantities between zero and six except equals. Although zero is maybe a special type of quantity, we called comparisons of zero each with one to six food items a test for food quantity. In contrast to Boysen et al. (1996), we did not drop out any array. Instead, each array was presented once per session that therefore consisted of 21 trials. The left–right position of the larger quantities was counterbalanced across trials with the constraint that the larger food quantity was never on the same side more than two times in a row. Between sessions, the left–right position of the larger quantity was alternated systematically for each array. Furthermore, the order of the 21 arrays per session was randomized

between all sessions. All subjects were presented with the same predetermined randomized order of arrays throughout testing. We administered 20 sessions to compare the results to the Boysen et al. (1996, 2001) studies, and to the results of the last study.

Procedure

All subjects were tested individually in an observation room. The experimenter (Jana Uher) baited the two dishes in the middle of the table behind the opaque barrier. When the subject was looking in her direction, she removed the barrier and strictly looked at the lower middle of the Plexiglas panel. Once the barrier was removed, the two bowls were moved outward from the center of the table. When the dishes reached their final position behind the two holes in the Plexiglas (but still out of reach for the ape) the experimenter immediately took her hands simultaneously back to her lap, and placed both thumbs close to another at the middle of her side of the table. She adopted this position at least shortly in every trial independent of the speed of the ape's choice.

The apes could select one of the two dishes by reaching with their fingers through the panel's holes next to the two dishes. If they reached through both holes simultaneously, the experimenter did not leave her position and only encouraged them verbally to choose one side only ("No, only one."). In the rare cases where apes insisted on a double choice, the dishes were pulled back for a moment and thereafter replaced following the same procedure described above. Furthermore, it was always made sure that the apes did not place their finger in one of the holes before a trial started. When the ape indicated his choice clearly, the experimenter looked up, pulled back the chosen dish and returned its content to the supply cup in full view of the subject. Thereafter, the nonchosen dish was pushed forward to the subject and its content was given to the ape. Thus, the apes were rewarded with the content of the dish they did not choose. If the nonchosen dish was

empty, it was equally pushed forward to the ape and shortly lifted vertically so that the ape could touch it through the hole. This sequential procedure is different from the simultaneous procedure in the last study, where the chosen dish was pulled back and the nonchosen dish was pushed forward simultaneously in a single movement.

Special attention was placed on preventing unintended cues that could have served a Clever Hans effect. After removing the barrier, the experimenter no longer looked at the ape but instead looked strictly down at the bottom of the Plexiglass. She only perceived in her visual field on which side the ape was sticking a finger through the Plexiglass panel, but not how the ape was behaving otherwise. To further prevent unintended cues by body posture, the experimenter always placed both thumbs next to another in the middle of her side of the table. Gaze aversion, looking down at the middle of the table, and placing the hands underneath the table except for the thumbs were used to prevent unintended cues. That some of the experienced apes failed in this task whereas some naïve apes mastered it paired with the present findings on the impact of quantity ratios on performance show that Clever Hans is very unlikely to account for the results.

Each session was administered on a different day. If subjects did not respond after 5 minutes, the session ended and was completed on another day. All sessions were videotaped for detailed coding. A second coder (Josep Call) scored 18% of the sessions to assess intercoder reliability. Intercoder reliability was excellent ($\kappa = .97$).

Results

Comparing Experienced and Naïve Subjects

Figure 1 presents the percentage of correct trials in the first and last sessions for naïve and experienced subjects. A mixed model analysis of variance (ANOVA) on the percentage of correct trials with experience as between-subjects factor, and session as within-subject factor indicated that experienced subjects performed significantly better than naïve subjects, $F(1, 13) = 7.90, p < .05, \eta^2 = 0.38$. In addition, subjects performed better in the last session compared to the first session, $F(1, 13) = 37.92, p < .001, \eta^2 =$

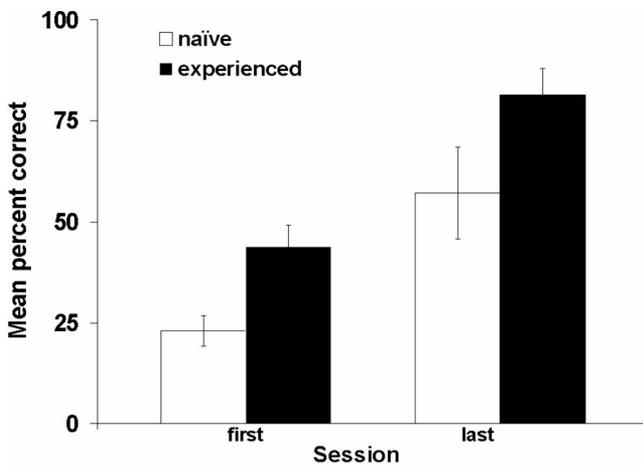


Figure 1. Mean percentage of correct trials (\pm SEM) in the first and last sessions for naïve and experienced subjects.

0.75, independently of experience, $F(1, 13) = 0.09, p > .70, \eta^2 = 0.007$. Experienced subjects performed at chance levels in the first session $t(8) = 1.16, p > .20$, and above chance levels in the last session, $t(8) = 4.94, p < .001$. In contrast, naïve subjects performed below chance levels in the first session $t(5) = 7.15, p < .001$, which means that they significantly preferred the larger quantity, and at chance levels in the last session, $t(5) = 0.63, p > .50$.

Although both groups of subjects improved their performance, they did it at different rates. Figure 2 presents the mean percentage of correct responses for each session as a function of experience. Experienced subjects followed a logarithmic progression ($\hat{y} = 0.1133 * \ln(\text{session \#}) + 0.5187, R^2 = 0.88, p < .001$), whereas naïve subjects followed a linear progression ($\hat{y} = 0.0226 * (\text{session \#}) + 0.1910, R^2 = 0.92, p < .001$). This means that within the first four sessions experienced subjects had quadrupled the improvement shown by naïve subjects (from 44% to 70% compared to 23% to 29%). In fact, experienced subjects had already reached above chance performance in the fourth session, $t(8) = 2.51, p < .05$.

Comparing the acquisition curves in the current study (see Figure 2) to those obtained by individuals in the Vlamings et al. (2006) study is additionally informative. Figure 3 presents the acquisition curves for the subjects that passed or failed the reversed reward contingency task in the Vlamings et al. study. Like the naïve subjects in the current study, both groups of subjects showed linear acquisition curves that differed in their slope (passers: $\hat{y} = 3.1056 * (\text{session \#}) + 0.8583, R^2 = 0.97, p < .001$; failers: $\hat{y} = 1.2852 * (\text{session \#}) + 5.0404, R^2 = 0.92, p < .001$). Note that although the subjects that passed the Vlamings et al. study were the experienced subjects in the current study, their acquisition curves were quite different, thus reinforcing the idea that experienced subjects had retained some knowledge of the task.

Individual analyses indicated that two of the six naïve subjects (one chimpanzee and one bonobo) reached a performance above chance levels (Binomial test: $p < .05$) in the 12th and 16th session, respectively. All experienced subjects except one orangutan and one gorilla reached a performance above chance levels (Binomial test: $p < .05$) between the second and the sixth session ($Mdn = 3$ sessions).

Comparing Pairs of Quantities

Next we tested whether experienced subjects found familiar pairs (4:0, 4:1) easier than unfamiliar pairs. Figure 4 presents the percent of correct trials for each pair of quantities. Although pairs varied in their difficulty, familiar pairs appeared no different from unfamiliar pairs. We refined this analysis by only considering the first three sessions for each of the two familiar pairs and compared them to two closely related unfamiliar pairs. Thus, we compared the performance of subjects in the 4:0 pair to the 3:0 and 5:0 pairs. There were no significant differences among pairs, $F(2, 16) = 0.57, p > .50, \eta^2 = 0.07$. Similarly, there were no differences among pairs if only the first session was considered (Cochran $Q = 2.8, df = 2, p > .20, N = 8$). Comparing the performance of subjects in the 4:1 pair to the 3:1 and 5:1 pairs produced analogous results. We found no significant differences among pairs in the first three sessions combined, $F(2, 16) = 0.12, p >$

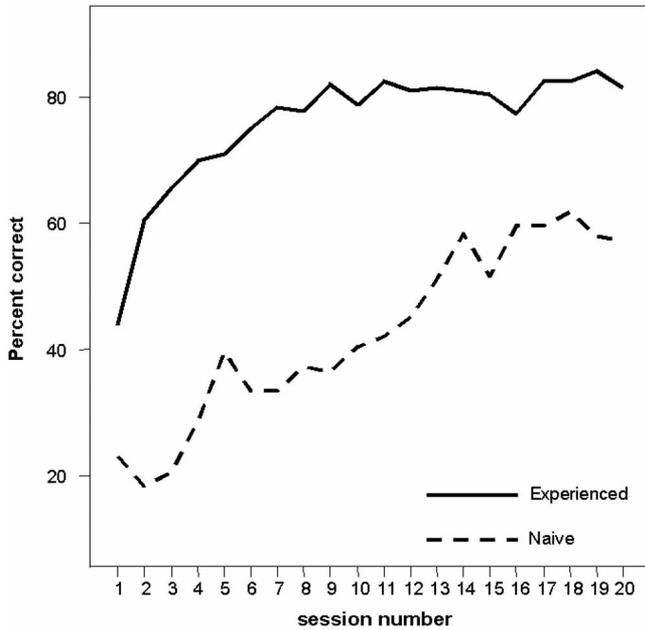


Figure 2. Mean percentage of correct trials for each session as a function of previous experience with the task.

.80, $\eta^2 = 0.02$, or the first session alone (Cochran $Q = 0.33$, $df = 2$, $p > .80$, $N = 8$).

Factors Influencing Success

Figure 4 indicated that pairs varied in the level of difficulty. We analyzed what factors may explain such a variability for experi-

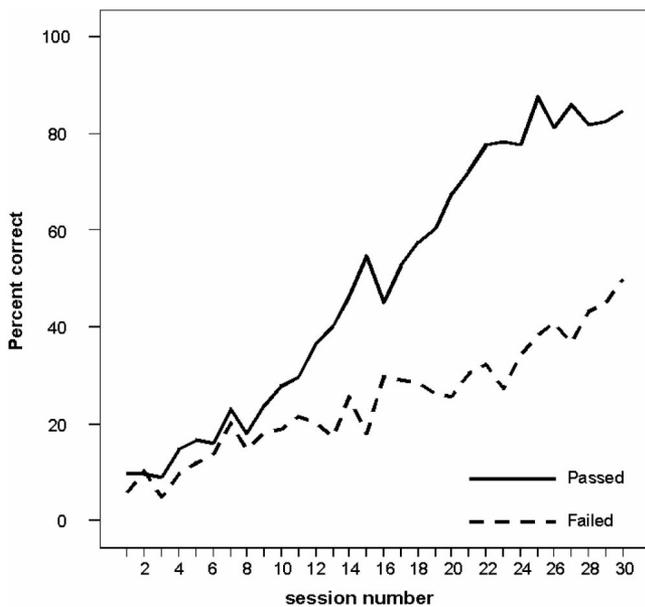


Figure 3. Mean percentage of correct trials for each session in the Vlamings et al. (2006) study as a function of final performance. The successful group participated as the experienced subjects in the present study.

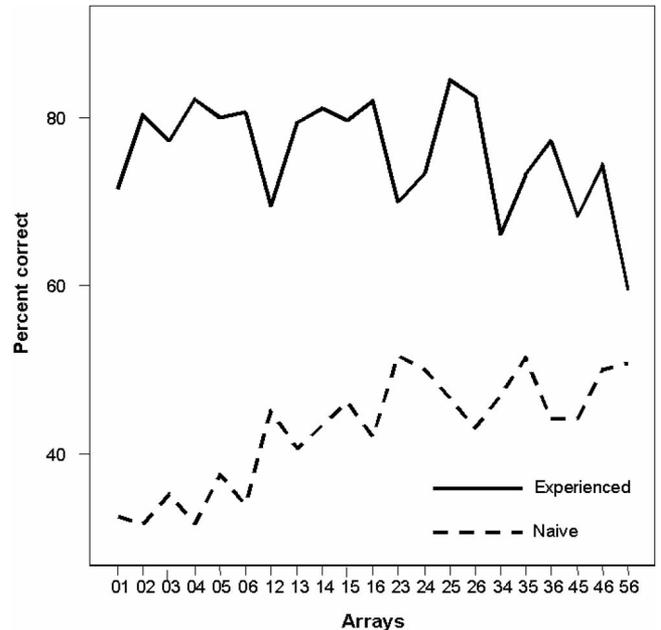


Figure 4. Mean percentage of correct trials for each pair of quantities as a function of previous experience with the task.

enced and naïve subjects. We decided to treat these two groups separately because their acquisition curves were also different (see Figure 2). For naïve subjects, a multiple-stepwise regression with ratio (smaller over larger quantity), disparity between quantities and total quantities as independent variables, and the percent of correct trials as dependent variable revealed that a model composed by the ratio and square of the ratio produced the best fit, $R = 0.916$, $F(2, 18) = 46.63$, $p < .001$. Subjects' performance followed an inverted U-shaped curve that was a function of the ratio and the square of the ratio ($\hat{y} = -0.313 * (\text{ratio}^2) + 0.428 * \text{ratio} + 0.341$). Ratio alone accounted for 74% of the variance in the model with ratio^2 explaining an additional 10%. In general, naïve subjects found pairs with smaller ratios harder than those with larger ratios (Figure 5a).

For experienced subjects, a multiple regression with ratio, disparity, and total quantities as independent variables and the percentage of correct trials as dependent variable revealed that a model composed by the square of the ratio, the total quantity, and the square of the total quantity produced the best fit, $R = 0.925$, $F(3, 17) = 33.69$, $p < .001$. Subjects' performance followed an inverted U-shaped curve that was a function of the square of the ratio and the total quantity ($\hat{y} = -0.294 * (\text{ratio}^2) - 0.002 * (\text{total}^2) + 0.034 * (\text{total}) + 0.701$). Ratio^2 alone accounted for 68% of the variance and total quantity and total quantity² explained an additional 12% and 6%, respectively. In general, experienced subjects found pairs with smaller ratios easier than those with larger ratios (Figure 5b).

Comparison With the Previous Study

Experienced subjects reached above chance performance in the current study significantly faster than when they faced the reversed reward contingency task for the first time, $t(6) = 8.72$, $p < .001$

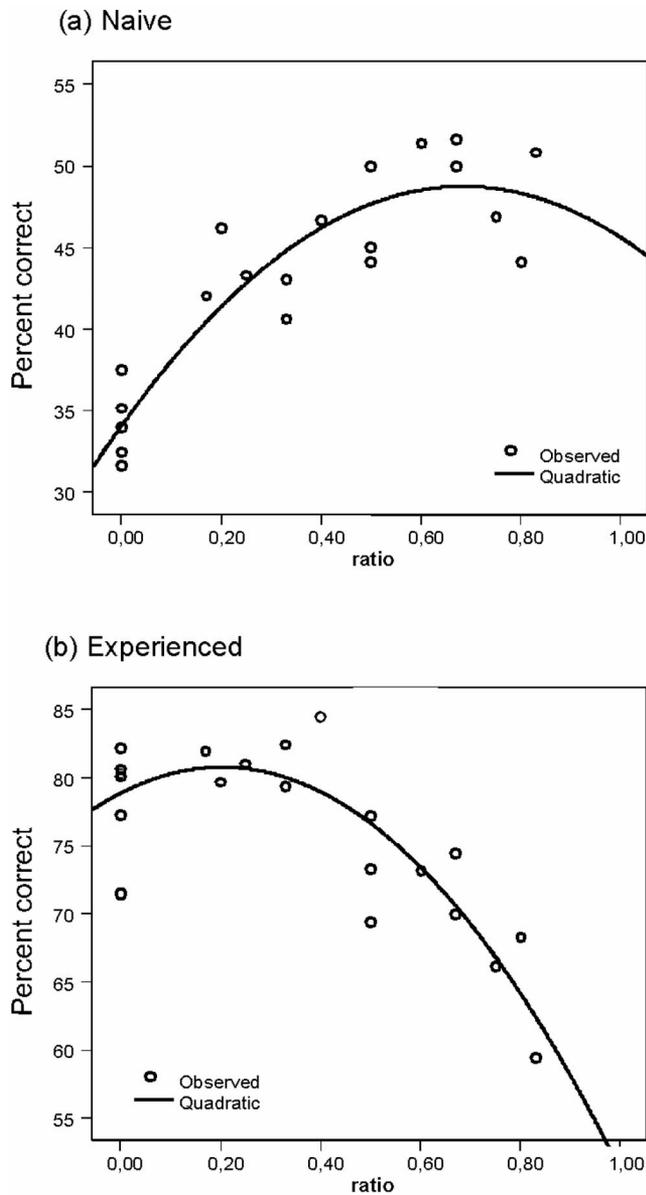


Figure 5. Mean percentage of correct trials as a function of the ratio (i.e., smaller quantity over larger quantity) between quantities for (a) naïve and (b) experienced subjects.

(number of trials current study: $M = 69.0$, $SEM = 11.9$; Vlamings et al., 2006, $M = 306.3$, $SEM = 25.0$). In contrast, naïve subjects from both studies performed in similar ways. In the previous study (Vlamings et al., 2006, Table 4), 6 out of 18 subjects performed above chance in both visible conditions, which are the most comparable to the current study. This is exactly the same proportion of success found in the naïve subjects of the current study (Fisher's exact test: $p = 1.00$). Moreover, the two successful naïve subjects in the current study reached the above chance performance in the same number of trials as those in the previous study, $t(10) = 0.38$, $p > .70$ (current study: $M = 294.0$, $SEM = 42.0$, $N = 2$; Vlamings et al., 2006, $M = 323.2$, $SEM = 32.5$, $N = 10$).

Discussion

Two naïve apes (1 chimpanzee and 1 bonobo) out of six were able to master the reversed reward contingency task using multiple-visible quantities from zero to six items and a reward scheme according to which the subject received the nonchosen quantity in all cases. We did not increase costs for the subjects by giving no reward at all when the larger quantity was chosen or by handing it over to a conspecific competitor. This means that the successful performance reported in the Vlamings et al. (2006) study cannot be solely attributed to the use of only two pairs of quantities or the concurrent use of colored containers representing quantities. In fact, the percentage of successful subjects with visible quantities was identical in both studies. In addition, subjects of the four great ape species that had mastered the task in that study showed evidence of transfer to multiple quantities in the current study. Such transfer was evident in the superior performance (including the first session) of experienced subjects compared to naïve ones. Experienced subjects also showed a steeper improvement in performance during testing than naïve subjects—an improvement that they had not shown themselves as naïve subjects in the previous study. This is a striking example of long-term procedural memory in a context in which one can easily argue that these apes should have “regressed” to the more natural response bias toward the larger array.

Experienced apes transferred their ability to arrays containing different quantities presented in the original study after 6 to 19 months had elapsed since the completion of the initial testing. This finding mirrors those reported in lemurs, squirrel monkeys, sea lions, and mangabeys that were shown to retain the reversed reward rule for 3 to 18 months (Albiach-Serrano et al., 2007; Anderson et al., 2004; Genty et al., 2004; Genty & Roeder, 2006). It also shows that subjects had learned a reversed reward contingency rule, not simply how to respond to two particular pairs of quantities. In fact, experienced subjects showed no advantage for familiar compared to unfamiliar pairs, including those that resembled the original pairs. In contrast, Boysen and Berntson (1995) found no transfer between Arabic numerals and pairs of quantities in an ABAB design. They reported a consistently poor performance with candy arrays but immediate improvement on the introduction of Arabic numerals followed by a sharp deterioration in performance when candies were used again. Moreover, the apes in the present study reached performance levels between 85% and 98% correct, whereas chimpanzees in other studies (not counting those included in Vlamings et al., 2006) remained at chance levels when presented with actual quantities (not Arabic numerals) except for one male that outperformed the rest of the group (Boysen et al., 2001).

The use of multiple-quantity combinations permitted us to analyze the factors influencing success. Ratio between quantities was the variable that explained most of the variance both for naïve (74%) and experienced individuals (68%), thus confirming the results from previous studies, even though some studies reported a much lower percentage of explained variation (e.g., 41%, Boysen et al., 2001). The influence of the ratio was different for naïve and experienced subjects. The greater the ratio the harder the task became for naïve subjects. In contrast, experienced subjects showed the opposite pattern—the greater the ratio the easier the task became. This apparent discrepancy helps us understand the

different factors that were influencing the subjects' choices in each group. We can assume that both groups had equivalent discriminative abilities—in fact, the data comparing their initial performance in this task support this assumption. This means that both groups discriminated pairs of quantities with smaller ratios better than quantities with larger ratios. The crucial difference between groups is that whereas naïve subjects had difficulties inhibiting their choice for the larger quantity, experienced individuals did not. This explains why for smaller ratios naïve individuals were well below 50% correct and experienced ones were well above 50%, yet both groups tended to converge toward 50% for larger ratios. Thus, while the curve for experienced subjects was equivalent to that observed in quantity discrimination task (e.g., Beran, 2001; Call, 2000; Hanus & Call, 2007; Huntley-Fenner & Cannon, 2000), the curve for naïve individuals is the opposite because they are not capable of inhibiting the selection for the larger quantity in the pair.

Other factors such as the specific reward delivery procedure appeared not to affect performance. In the current study, unlike Vlamings et al. (2006), but like Boysen et al. (1996, 2001), we used a successive procedure (as opposed to simultaneous) that consisted of removing the chosen contents first and only then giving the nonchosen reward to the subject. However, experienced apes showed no problem adjusting to this new procedure and more important, some naïve subjects also mastered the task. We propose that the major reason for differences in performance across studies is due to the existence of striking individual differences. In most studies, some individuals mastered the task and some failed. Among those who failed in the present and previous study, some developed a strong preference for one side that due to the randomization permitted them to obtain performances at the chance level. Other individuals, by contrast, continued to have a stronger preference for the larger quantity regardless of the side on which it was presented; their performances remained below chance levels. These differences may reflect differences in the strategies individual apes develop in trying to master the task and in their ability to overcome their strong prepotent response to reach for the larger quantity.

Larger samples of naïve apes in particular are needed to add further empirical evidence to the present findings. This especially applies to the bonobos, gorillas, and orangutans of which we had either only one or no naïve subject available for the present study. Empirical evidence to the importance of individual differences is accumulating across species (Anderson et al., 2000; Boysen et al., 2001; Murray et al., 2005; Vlamings et al., 2006). Larger samples will help clarify the importance of individual differences relative to species differences in inhibitory control (Bell, 1998). Therefore, future studies are important to understand the evolution of interspecific differences in problem solving abilities (Kralik et al., 2002).

References

- Albiach-Serrano, A., Guillén-Salazar, F., & Call, J. (2007). Mangabeys (*Cercocebus torquatus lunulatus*) solve the reverse contingency task without a modified procedure. *Animal Cognition*, *10*, 387–396.
- 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.
- 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.
- Beran, M. J. (2001). Summation and numerosness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *115*, 181–191.
- 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.
- Cheney, D. L., & Seyfarth, R. M. (1985). Vervet monkey alarm calls: Manipulation through shared information? *Behaviour*, *94*, 150–166.
- Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the human brain*. New York: Norton.
- 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: Academy of Sciences.
- 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.
- Genty, E., & Roeder, J. J. (2006). Self-control: Why should sea lions (*Zalophus californianus*) perform better than primates? *Animal Behaviour*, *72*, 1241–1247.
- Hanus, D., & Call, J. (2007). Discrete quantity judgements in the great apes: The effect of presenting whole sets vs. item-by-item. *Journal of Comparative Psychology*, *121*, 241–249.
- Hauser, M. D. (1993). Rhesus monkey (*Macaca mulatta*) copulation calls: Honest signals for female choice? *Proceedings of the Royal Society of London B*, *254*, 93–96.
- Huntley-Fenner, G., & Cannon, E. (2000). Preschoolers' magnitude comparisons are mediated by a preverbal analog mechanism. *Psychological Science*, *11*, 147–152.
- 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. (1974). A group of young chimpanzees in a one-acre field:

- Leadership and communication. In A. M. Schrier & F. Stollnitz (Eds.), *Behavior of nonhuman primates* (pp. 83–153). New York: Academic.
- 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.
- 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*, *15*, 385–391.
- Silberberg, A., & Fujita, K. (1996). Pointing at smaller food amounts in an analogue of Boysen & Berntson's (1995) procedure. *Journal of the Experimental Analysis of Behaviour*, *66*, 143–147.
- Vlamings, P. H. J. M., Uher, J., & Call, J. (2006). How the great apes (*Pan troglodytes*, *Pongo pygmaeus*, *Pan paniscus*, and *Gorilla gorilla*) perform on the reversed contingency task: The effects of food quantity and food visibility. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*, 60–70.

Received May 5, 2007

Revision received December 13, 2007

Accepted December 21, 2007 ■