

Gravity and Solidity in Four Great Ape Species (*Gorilla gorilla*, *Pongo pygmaeus*, *Pan troglodytes*, *Pan paniscus*): Vertical and Horizontal Variations of the Table Task

Trix Cacchione
University of Zürich

Josep Call
Max Planck Institute for Evolutionary Anthropology

Robert Zingg
Zoo Zürich

Three experiments modeled after infant studies were run on four great ape species (*Gorilla gorilla*, *Pongo pygmaeus*, *Pan troglodytes*, *Pan paniscus*) to investigate their reasoning about solidity and gravity constraints. The aims were: (a) to find out if great apes are subject to gravity biased search or display sensitivity for object solidity, (b) to check for species differences, and (c) to assess if a gravity hypothesis or more parsimonious explanations best account for failures observed. Results indicate that great apes, unlike monkeys, show no reliable gravity bias, that ape species slightly differ in terms of their performance, and that the errors made are best explained by a gravity account.

Keywords: invisible displacement, object solidity, gravity error, naïve physics

Various constraints, such as gravity, inertia, or solidity, act upon a moving object, shaping its pathway through time and space. To locate an invisibly displaced object, an infant must have some knowledge of how these constraints interact to determine the future trajectory of the object moving out of sight (Baillargeon, 2002, 2004; Berthier, DeBlois, Poirer, Novak, & Clifton, 2000; Butler, Berthier, & Clifton, 2002; Hood, 1995, 1998; Hood, Carey, & Prasada, 2000; Hood, Santos, & Fieselman, 2000; Kim & Spelke, 1992, 1999; Hood, Cole-Davies, & Dias, 2003; Mash, Keen, & Berthier, 2003; Spelke, Breinlinger, Macomber, & Jacobson, 1992).

Using a looking time methodology, Spelke, Breinlinger, Macomber, and Jacobson (1992) set out to investigate at what age infants understand how solidity and gravity constrain the path of an invisibly moving object. For this purpose, they invented a task that is generally referred to as the table task and was later adopted by various researchers (e.g., Berthier et al., 2000; Hauser, 2001; Hood et al., 2000; Hood, Cole-Davies, & Dias, 2003; Hood, Santos, & Fieselman, 2000). In a solidity version of the task, infants were presented with a solid shelf placed above a stage floor. A screen was set in place and an object dropped behind it. The removal of the screen revealed the object either resting on top of the shelf (possible test event) or under the shelf on the stage floor (impossible test event because object solidity would prevent

the object from traveling through the shelf). The idea is that a preference to look at the impossible test event suggests that infants detect the physical anomaly of the event and have, therefore, some knowledge of solidity constraints. To test for gravity, Spelke and colleagues presented the same task, but this time no shelf was in place: The impossible test event showed the object floating in midair, while it rested on the stage floor in the possible test event. Again, if infants have some knowledge of gravity constraints they should look longer at the unsupported yet stationary object. Using these tasks, Spelke et al. (1992) found no evidence that 3 to 4-month-old infants expected the unsupported object to fall. At the same age, however, infants' reasoning seemed to be in accord with the core principles of continuity (an object moves on exactly one connected path over space and time) and solidity (two objects cannot occupy the same place at the same time, i.e., cannot move through each other), as a horizontal version of the task suggested. Spelke et al. (1992) argued that the negative results support the "core knowledge thesis" and the assumption that, while knowledge about object solidity is innate, knowledge about gravity develops slowly with increasing experience.

A completely different picture emerges when young children are tested with tasks involving overt action. Despite the precocious hints of knowledge reported by Spelke and colleagues, Hood (1995) did not find older children's search behavior to be in accord with solidity. He presented 2 to 3-year-olds with a tubes task in which the goal was to find a ball dropped down one of a set of three interwoven tubes landing in one of three hiding places. Hood identified an important phenomenon labeled "gravity bias." Two-year-old children failed to search the object in the box attached to the tube. However, errors were not at random but occurred significantly more often at the hiding place directly underneath the release point, suggesting some knowledge of gravity to be present in 2-year-olds. Hood, Santos, and Fieselman (2000) did not find a

Trix Cacchione, Department of Psychology, Cognitive and Developmental Psychology, University of Zürich, Switzerland; Josep Call, Max Planck Institute for Evolutionary Anthropology, Developmental and Comparative Psychology, Leipzig, Germany; Robert Zingg, Zoo Zürich, Switzerland.

Correspondence concerning this article should be addressed to Trix Cacchione, Department of Psychology, University of Zürich, Binzmühlestrasse 14/21, 8050 Zürich. E-mail: t.cacchione@psychologie.uzh.ch

similar bias on the horizontal plane, even though children's overall search performance was not better. Additionally, Hood (1998) found 2-year-olds responses were significantly less biased when he presented the falling motion in reverse so that the object traveled through the tube in upward direction. Hood interpreted the bias as stemming from a naïve folk theory of gravity, which says that all unsupported objects fall in a straight line. The *gravity hypothesis* proposes that children develop a sensitivity for gravity by detecting statistical regularities in the myriad of gravity events that they witness in their everyday lives. This sensitivity leads to specific expectations about the future behavior of (unseen) objects and culminates in a response bias that may lead to the production of incorrect responses. Three-year-olds overcome the bias because they acquire the ability to inhibit unwarranted gravity responses. That is, the bias is not abandoned but remains dormant and may reappear under certain circumstances. Hood, Wilson, and Dyson (2006) demonstrated that 4-year-old children show the bias again when the cognitive load of the tubes task is increased by dropping two objects simultaneously. In specific situations, the gravity bias even appears in adults (see Bliss, Ogborn, & Whitelock, 1989; Kaiser, Profitt, & McCloskey, 1985; Krist, 2000).

However, several authors have challenged the gravity hypothesis on theoretical and empirical grounds. Several studies on horizontal trajectories found no evidence of performance improvements (Berthier et al., 2000; Frye, Zelazo, & Palfai, 1995; Hood et al., 2000; Mash et al., 2003). In one case, similar perseveration in horizontal plane displacements was even reported (Frye et al., 1995). Moreover, it is unclear whether the findings from upward trials really can be interpreted as supporting the gravity hypothesis. Hood (1998) presented both the upward trials and downward trials as films on TV monitor and found that children were less biased than in the original search task on the downward trials, although still to a greater extent than in the upward version. This weakening of the bias can, therefore, be explained by the different mode of presentation. In fact, Southgate (2005) found the error also in upward direction, if real events were presented.

A possible problem of the tubes task may be that it confounds multiple principles, such as solidity, gravity, and straight trajectories. To directly compare performance, Hood, Carey, and Prasada (2000) tested knowledge of solidity and support in four search tasks adapted from looking time studies (Spelke et al., 1992). In one of the tasks, children saw a toy being dropped behind a screen and were then asked to search for it. In a solidity condition, a shelf was inserted above the stage floor; in a support condition, the shelf was removed. In both conditions, search behavior was recorded (search in the upper or in the lower location). Results showed that until 2½ years of age, neither solidity nor support guides search behavior. No systematic search error was found in this task context, instead children tended to search for the object where they saw it during familiarization. Therefore, the results are in conflict not only between looking and search tasks but also within different types of search tasks. In the table task, no gravity error appears, even though search performance is not correct.

The knowledge about physical object properties has also been addressed using a comparative perspective. Various studies have investigated reasoning about gravity and solidity constraints in nonhuman primates (Cacchione & Krist, 2004; Hauser, 2001; Hauser, Williams, Kralik, & Moskovitz, 2001; Hood, Hauser, Anderson, & Santos, 1999; Southgate & Gómez, 2006; Tomonaga,

Imura, Mizuno, & Tanaka, 2007) and dogs (Osthaus, Slater, & Lea, 2003). Addressing the question of object knowledge from a comparative standpoint can allow us to reconstruct the evolution of this domain as well as the potential effects of certain ontogenetic trajectories. Hood et al. (1999) presented adult cotton-top tamarins (*Saguinus oedipus oedipus*) with a version of the original tubes task. Most tamarins failed to solve the task, but as with children search was not at random: they too seem predisposed to search erroneously in the gravitational direction, that is, in the box directly underneath the release point. Osthaus, Slater, and Lea (2003) ran the tubes task on dogs (*Canis lupus familiaris*). Dogs, like human infants and tamarins, were biased to search the box in the direction of gravity, but unlike monkeys they learned to search in the correct location, thus overcoming the bias. This finding suggests that the liability to gravity biases might differ across species.

Recently, Tomonaga, Imura, Mizuno, and Tanaka (2007) confronted young and adult chimpanzees with a modified tubes task in which subjects had to predict the landing location of objects before they were dropped into two either parallel or crossed tubes. Note that this was a prediction task (i.e., subject had to choose the location before the object was dropped), which is not fully comparable to the original task. The authors found the chimpanzees' predictions were biased in the direction of gravity. Neither previous experience with tubes nor learning over multiple trials improved their performance. These results are in contrast with our own findings (Cacchione & Call, 2008). We confronted orangutans, bonobos, gorillas, and chimpanzees with versions of the original tubes task. Despite the fact that the gravity box was initially more often selected than the two other locations and more gravity than nongravity errors occurred, the preference for the gravity location did not reach statistical significance and was abandoned in relatively short time. However, the gravity error was not completely absent in apes. When they were confronted with a situation where no logical solution could be inferred (i.e., when presented with a painted two-dimensional line instead of a real three-dimensional tube), they reverted to gravity responses as a default answer. This might suggest that, as three-year-old children, apes can overcome an initial bias due to the recruitment of inhibitory control. This is remarkable because, in contrast to human children, monkeys apparently never acquire the inhibitory skills to suppress the gravity bias (Gómez, 2005; Hauser, 2003).

However, some authors did not interpret the behavior observed in the above mentioned studies as being caused by a gravity bias, but as being part of a more general search pattern based on extrapolating straight object trajectories. Thus, the proximity hypothesis states that children (and monkeys) search in the closest location to the position where the object was last seen (Hood, 1995, 1998; Southgate & Gómez, 2006), which in the case of the tubes task corresponds to the box directly located underneath the release point. That a similar bias was found in upward (Southgate, 2005) and in horizontal directions (Frye et al., 1995) supports this account, although some findings on the horizontal plane contradict it (Hauser, 2001).

Hauser (2001) investigated the presence of gravity biased search in the table task. He presented rhesus macaques (*Macaca mulatta*) with a table and two containers, one on top of the table and one under the table. The table was subsequently covered by an occluder and a food reward was dropped behind it. After the removal of the occluder, the monkey was allowed to search for the reward.

In contrast to human children, monkeys showed a strong search bias toward the box located below the table. Note that these findings contradict the proximity hypothesis, which predicts that subjects should search the top location. Furthermore, Santos and colleagues reported dissociations between looking and action responses in rhesus macaques (*Macaca mulatta*) and cotton-top tamarins (*Saguinus oedipus oedipus*) equal to those found in human infants (Santos & Hauser, 2002; Santos, Seelig, & Hauser, 2006). For example, rhesus macaques looked longer at an object apparently falling through a solid table even though they searched for it under the table in an action context (Santos & Hauser, 2002). Again, this suggests that subjects might not simply rely on a proximity search rule. Finally, in their modified tubes task Tomonaga et al. (2007) found some evidence that chimpanzees did not rely on spatial proximity.

Two additional findings appeared to reinforce the notion of a gravity-based search bias. First, shifting the goal box away from the vertical falling trajectory reliably improved the search performance whereas a slight deviation did not (Hauser, 2001; Southgate & Gómez, 2006, found similar results, but offer a different interpretation, see below). Second, monkeys presented with a horizontal version of the task performed well. These findings are further supported by results from a horizontal version of the tubes task (Hauser et al., 2001), where performance improvement in case of horizontal displacement was also found. However, Southgate and Gómez (2006) argued that searching at the bottom location (i.e., under the table-shelf) may not reflect a gravity bias but a preference for foraging in sheltered locations (*shelter hypothesis*) in order to avoid food competition from conspecifics (see also Karin-D'Arcy & Povinelli, 2002).

In sum, children and monkeys show important search biases in gravity/solidity tasks, although children can overcome them after a certain age. Regarding the behavior of our closest living relatives, the great apes, the present findings are contradicting. One important piece of the puzzle that is missing is the behavior of great apes in the table task. These data are crucial to be able to make inferences about the evolution of the cognitive processes involved in judging solidity and gravity. Thus, the first goal of this study was to investigate the performance of all four great apes in the table task. We chose this task because it is less complex than the tubes task and directly modeled after the looking tasks used in infancy (Spelke et al., 1992). Would the behavior of apes on this task mirror that of rhesus macaques (bottom location bias), 2½-year-old children (no bias), or older children and adults (upper location search)?

Our second goal was to investigate how experience affected subjects' responses. Hauser (2003) argued that tamarins (in contrast to human infants where the bias is a transitory developmental phenomenon) were unable to suppress responses based on gravity bias, but instead they persisted on their straight down belief even in the face of massive contrary evidence. This may be particularly true for the tubes task in which tamarins persisted on their errors despite receiving feedback about the correct food location. Rhesus macaques in the table task only received a single trial, so it is unknown whether they would be able to overcome their gravity bias. Data on dogs (Osthaus et al., 2003) suggest that the bias may not be as strong as it has been stressed, at least not in all species. Our own data suggests that apes might in fact be able to overcome an initial gravity bias (Cacchione & Call, 2008). Thus, we wanted

to test whether subjects may alter their initial responses after multiple trials.

Our final goal was to assess which of the hypotheses that have been postulated fit our data best. To decide between the various alternatives, we conducted three experiments that allowed us to contrast the different hypotheses by making a set of predictions. Experiment 1 consisted of the original vertical table task in which we varied the length of the falling trajectory (short vs. long) before the object disappeared behind the screen. Both the gravity and the shelter hypotheses predict searching at the bottom location whereas the proximity and the solidity hypotheses predict searching at the top. Moreover, the gravity hypothesis predicts a larger bias in the long object trajectory compared to the short trajectory, while all other hypotheses make no prediction in this respect (see explanation in the introduction of Experiment 1).

Experiment 2 consisted of a horizontal object displacement with subjects witnessing either a long or a short trajectory before the object disappeared behind the screen. Both the proximity and solidity hypotheses predict searching in the near location whereas the shelter hypothesis makes no specific prediction. Although the gravity hypothesis also makes no specific prediction about the precise search location (particularly, it cannot explain the nature of potential failures with horizontal trajectories), it nevertheless predicts that subjects should perform better in the horizontal compared to the vertical task if the source of errors is gravity related reasoning (Experiment 1 & 3). Additionally, it predicts that the length of the object trajectory should not influence performance on horizontal trials.

Experiment 3 investigated whether subjects were affected by the degree of misalignment (large vs. slight) between the object's release point and the target location in vertical displacements. Like Experiment 1, both the gravity and the shelter hypotheses predict searching at the bottom location whereas the proximity and the solidity hypotheses predict searching at the top. Additionally, the gravity hypothesis predicts that subjects should be more likely to search the top location after a large misalignment compared to a slight misalignment. All other hypotheses make no prediction in this respect.

General Method

Participants

We tested 10 gorillas, 11 orangutans, 6 chimpanzees, and 5 bonobos (*Gorilla Gorilla*, *Pongo Pygmaeus*, *Pan Troglodytes*, *Pan Paniscus*; see Table 1). Five gorillas and three orangutans were housed in the Zürich Zoo, Switzerland. All other apes were housed at the Wolfgang Köhler Primate Research Center in Leipzig Zoo, Germany. The Leipzig group had a history of experience with various experiments concerning physical cognition. The Zurich group was completely naïve to experimental testing. Three additional orangutans and two additional gorillas from Zurich had to be excluded from the sample due to nonattendance.

Materials

The apparatus was placed on a table like plastic testing surface. It consisted of a gray rectangular four-legged plastic table (height 23.8 cm/ length 42.9 cm/ depth 33 cm) and two oblong blue plastic

Table 1
Species, Age, Sex, Birthplace, and Rearing History of the Apes Participating in the Experiments 1–3

Participant	Species	Age (years)	Sex	Birthplace	Rearing history
Azizi	Gorilla	5	m	Zurich	Mother
Bebe	Gorilla	25	f	Cameroon	Mother (wild born)
Binga	Gorilla	4	m	Zurich	Mother
Bonsenga	Gorilla	4	m	Zurich	Mother
Gorgo	Gorilla	24	m	Krefeld	Hand reared
N'Diki	Gorilla	27	f	Cameroon	Mother (wild born)
N'Gola	Gorilla	28	m	Jersey	Mother
N'Yokumi	Gorilla	4	f	Arnheim	Foster mother/Hand reared
Ruby	Gorilla	7	f	Arnheim	Hand reared
Viringika	Gorilla	9	f	Zurich	Mother
Bimbo	Orangutan	25	m	Duisburg	Hand reared
Dokana	Orangutan	16	f	Dresden	Mother
Dunja	Orangutan	32	f	Berlin	Mother
Oceh	Orangutan	17	f	Zurich	Mother
Padana	Orangutan	8	f	Leipzig	Mother
Pini	Orangutan	17	f	Leipzig	Mother
Salih	Orangutan	13	f	Zurich	Mother
Toba	Orangutan	11	f	Leipzig	Mother
Walter	Orangutan	16	m	Frankfurt a. M.	Mother
Xira	Orangutan	8	f	Zurich	Mother
Zora	Orangutan	15	f	München	Hand reared
Joey	Bonobo	23	m	Antwerpen	Hand reared
Kuno	Bonobo	9	m	Stuttgart	Hand reared
Limbuko	Bonobo	10	m	Stuttgart	Hand reared
Ulindi	Bonobo	12	f	Frankfurt a. M.	Mother
Yasa	Bonobo	8	f	Warwickshire	Mother
Alex	Chimpanzee	4	m	Plaisance du Touch	Hand reared
Alexandra	Chimpanzee	6	f	Rijswijk	Hand reared
Annette	Chimpanzee	6	f	Rijswijk	Hand reared
Fifi	Chimpanzee	12	f	Rijswijk	Mother
Jahaga	Chimpanzee	12	f	Rijswijk	Mother
Trudi	Chimpanzee	12	f	Rijswijk	Mother

cups (height 13.5 cm/ diameter 8.5 cm). Cups were filled with screwed up paper towels to prevent the reward from making a noise upon dropping into the cup. One cup was placed on the tabletop, the other beneath it. A plastic screen (43.5 cm × 68.4 cm) was used to block the subjects' visual access during the object dropping event. Grapes served as the dropped object for most subjects except for those that for dietary reasons had to be tested with banana slices.

Design

We administered two experiments (Exps. 1 & 3) on vertical displacement and one experiment (Exp. 2) on horizontal displacement. Each participant started with one of the vertical experiments, progressed to the horizontal experiment, and finished with the remaining vertical experiment. All three experiments consisted of two test conditions counterbalanced between subjects. Additionally, Experiment 2 and 3 included a control condition to detect side biases.

First, we administered two baseline conditions involving vertical displacements to screen out those subjects unable to master the basic test prerequisites. The *vertical baseline condition* served to exclude potential preferences for one of the two cup positions. The *vertical switch condition* assessed whether subjects were able to switch to a new location after finding the object repeatedly in the same location. For sake of simplicity, the detailed procedure and

results of the vertical baseline conditions are presented in Experiment 1. After successfully completing both baseline conditions, which all subjects did, half of the participants received Experiment 1 and the other half Experiment 3. We counterbalanced the order of administration of each experiment to control for potential carry-over effects between experiments. Next, all apes received two horizontal baseline conditions analogous to the vertical baselines before receiving the two test conditions of Experiment 2. Finally, those participants who were first run on Experiment 1 now passed on to Experiment 3 and vice versa.

Procedure

We tested subjects alone in an indoor room. Mothers with dependent offspring were tested in company of their offspring. Some of the Zurich gorillas could not be separated from the group; they were tested in groups of two to four individuals, in which each was in turn lured to the test location while a keeper kept the others occupied.

The ape sat behind a mesh panel. The apparatus was placed on top of a horizontal platform perpendicular to the mesh panel. The experimenter sat in front of the platform that could be moved toward or away the mesh panel. In Leipzig, the sliding mechanism consisted of a table mounted on rails whereas in Zurich consisted of a metal trolley. The experimenter pulled the table back to present the displacement events (dropping/rolling of object), then

she pushed it to the mesh to allow the ape to search for the food item by pointing to the cup location where it expected the food to be hidden. If the ape selected the correct cup, he received the food reward. If it pointed incorrectly, he was shown the empty cup and was allowed to choose again. This procedure was repeated until the ape located the food item or abandoned search.

Each of the two baseline conditions amounted to 10 trials. In the test conditions, the apes were tested until they reached a criterion: If they were correct on the first choice in four out of five trials, the session was terminated. Otherwise they received 10 trials.

Data Scoring and Analysis

We videotaped all trials and scored them live on coding sheets. We analyzed both first trial and overall performance on the various conditions comparing the performance of all great ape species. We further checked for potential species differences. Because of differential sample sizes we analyzed both, the data of the complete sample and the data of the five bonobos (the maximum number of bonobos available) with five each of the other ape species selected to match each other as much as possible (i.e., by closest age and environment). The restricted sample contained only apes of the Leipzig group (in fact all apes tested in Leipzig minus one chimpanzee and three oranges). We also assessed the occurrence of learning by comparing the performance in the 1st and the 4th trials, as this enabled us to include all apes in the sample. Recall that those subjects that responded correctly in 4 out of 5 trials did not receive a second round of tests. Thus, Trial 4 was the last trial that all apes completed. Finally, age analyses were done. We used Monte Carlo adjustment with a level of confidence 95% for all nonparametric tests to reduce the possibility of Type I error. Whenever possible we used the Pearson contingency coefficient "C" to assess effect sizes (Cohen, 1988). Next, we outline the procedure and results of Experiments 1 to 3 in sequence; an in-depth examination of test order effects follows Experiment 3.

Experiment 1: Length of the Vertical Trajectory

This experiment investigated whether great apes display a gravity bias and, if so, how persistently this influences search behavior across trials. Additionally, this experiment investigated whether witnessing a longer object fall exacerbated the possible bias. The impulse to search for a falling object straight down is an adaptive response acquired by vast experience with falling events. Thus, this response might be triggered when seeing an object fall. For example, Southgate, Gómez, Fox, and Meints (2005) found that the motion of a falling object triggered straight-down saccades in human children, which might cause the gravity responses. Further, the ability to suppress this reaction might be directly linked to the perception of the fall. It is possible that apes are more likely to inhibit a gravity response when they do not or only partially see the fall of the object (or are less likely to suppress it after witnessing a long object fall, respectively). Thus, this experiment consisted of the original vertical table task in which we varied the length of the falling trajectory (short vs. long) before the object disappeared behind the screen. Both the gravity and the shelter hypotheses predict searching at the bottom location, whereas the proximity and the solidity hypotheses predict searching at the top. Moreover, the gravity hypothesis predicts a larger bias (i.e., a greater propen-

sity to choose the lower container) in the long object trajectory compared to the short trajectory, while all other hypotheses make no prediction in this respect.

Procedure

Prior to the test, apes received two baseline conditions that served to a) familiarize participants with the apparatus, b) exclude baseline preferences for specific cup locations, and c) exclude inhibitory problems influencing the search behavior.

Vertical baseline condition. One cup was positioned on top of the table, the other beneath it. The experimenter held out a grape and in full view of the ape placed it in one of the cups. Then the apparatus was pushed against the mesh. In order to get the food reward, the ape had to correctly point out the cup containing it. If they selected correctly, the experimenter took out the food reward and gave it to the ape. If they were wrong, she picked up the selected cup and turned it to the ape, so it could see it was empty. Subjects received 10 trials with the reward placed randomly in both cups (five times in the top cup and five times in the bottom cup). If the search was correct in 9 out of 10 trials the participant moved on to the *vertical switch condition*. Otherwise, the baseline condition was repeated.

Vertical switch condition. The ape was presented with the same setup except that during the first five trials, the grape was placed in the same cup (the position was counterbalanced across apes). In the sixth trial, the location of the grape was changed, and in the last four trials it was again placed randomly in either location. If the ape chose correctly in the sixth trial (and could therefore suppress the tendency to select the location previously rewarded), it moved on to the vertical test conditions of either Experiment 1 or 3. If it failed, the *vertical switch condition* was repeated.

Upon passing both baseline conditions (something that all apes did), they received the following two conditions (see Figure 1) in counterbalanced order:

Vertical far drop condition. Again, one cup rested on top of the table while the other was placed beneath it (see Figure 1). The experimenter raised a screen and waved the grape 25 cm above the screen and the cups. If the ape tracked it, the experimenter released it. The screen was removed and the apparatus pushed to the mesh so that apes could indicate their choice. If they selected the correct upper cup, they received the grape. If they pointed incorrectly, they were shown the empty cup and were allowed to choose again. This procedure was repeated until the apes located the food item or abandoned search.

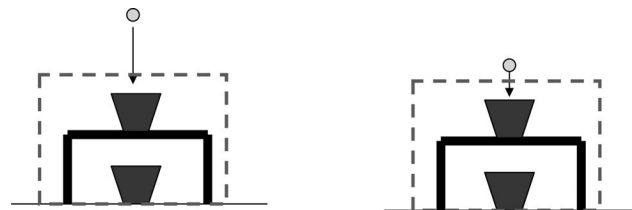


Figure 1. Set up for the test conditions in Experiment 1. Vertical far drop condition is depicted on the left side, the vertical near drop condition on the right side. The green dotted line indicates the screen.

Vertical near drop condition. This condition was identical to the *vertical far drop condition* except that the grape was dropped directly above the screen.

Results

First Exposure to the Task (Including Either Far or Near Drop Conditions)

Because the two test conditions were counterbalanced across subjects, we analyzed their choice in the first trial of their first exposure to the task (either far or near drop condition) to check for potential biases that might be washed out in the course of the task. Apes showed no bias for the bottom or top location in the first trial, $\chi^2(1, N = 32) = 1.125, p = .38, C = .18$, with 19 of 32 apes choosing the correct upper location. Although performance improved slightly in the fourth trial (21 of 32 subjects chose correctly), subjects still failed to select the top cup above chance levels, $\chi^2(1, N = 32) = 3.125, p = .10, C = .30$. Fifty-six percent of the apes selected the correct cup in at least four of the first five trials, $\chi^2(1, N = 32) = 0.500, p = .48, C = .12$, whereas 84% reached this criterion within 10 trials, $\chi^2(1, N = 32) = 15.125, p < .001, C = .57$.

The four ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, N = 32) = 3.097, p = .38$, or in their cup selection in the first trial, $\chi^2(3, N = 32) = 7.477, p = .06, C = .62$, or fourth trial, $\chi^2(3, N = 32) = 2.031, p = .60, C = .35$. However, African apes, represented by gorillas, chimpanzees, and bonobos, chose the top location more often than orangutans in the first trial, $\chi^2(1, N = 32) = 7.161, p < .05, C = .43$, but this difference disappeared by the fourth trial, $\chi^2(1, N = 32) = 0.912, p = .44, C = .17$. To equalize sample sizes across species we reran the analyses including only five (the maximum number of bonobos available) representatives of each species selected to match each other as much as possible. Ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, n = 20) = 7.166, p = .06$. However, they differed in their cup selection in the first trial, $\chi^2(3, n = 20) = 10.000, p < .05, C = .82$, but not in the fourth trial, $\chi^2(3, n = 20) = 4.762, p = .44, C = .62$. African apes, represented by gorillas, chimpanzees, and bonobos, chose the top location more often than orangutans in the first trial, $\chi^2(1, n = 20) = 10.000, p < .01, C = .58$, but this difference disappeared by the fourth trial, $\chi^2(1, n = 20) = 2.857, p = .09, C = .35$. Correlating age with percentage of correct choices revealed that younger apes did not

differ from older apes in their overall preference for the top container, Spearman correlation $r(32) = -.27, p = .13$.

Vertical Far Drop Condition

Table 2 presents the percentage of subjects that selected the top container in the first and fourth trials (far and near conditions) for each species. Overall, apes selected the top container in 70% of the trials, Wilcoxon's test: $z = 2.75, p < .01$, Table 2. Although this preference for the top container was not evident in the first trial, $\chi^2(1, N = 32) = 1.125, p = .37, C = .18$, it was apparent by the fourth trial as 23 of the 32 apes chose the top container, $\chi^2(1, N = 32) = 6.125, p < .05, C = .40$. Sixty-six percent of the apes selected the correct cup in at least four of the first five trials, $\chi^2(1, N = 32) = 3.125, p = .11, C = .30$, whereas 84% reached this criterion within 10 trials, $\chi^2(1, N = 32) = 15.125, p < .001, C = .57$.

The four ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, N = 32) = 0.760, p = .87$, or their preference in the first, $\chi^2(1, N = 32) = 1.485, p = .77, C = .30$, and fourth trial, $\chi^2(1, N = 32) = 1.434, p = .71, C = .29$. The same pattern appeared when comparing only five individuals of each species. The four ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, n = 20) = 1.893, p = .62$, or in their preference in the first, $\chi^2(3, n = 20) = 1.667, p = .96, C = .39$, or the fourth trial, $\chi^2(3, n = 20) = 2.933, p = .67, C = .50$. The overall preference for the top container did not differ across age, Spearman correlation $r(32) = -.17, p = .35$.

Vertical Near Drop Condition

Overall, apes selected the top container in 79% of the trials, Wilcoxon's test: $z = 3.80, p < .001$; see Table 2. Although this preference for the top container was not evident in the first trial (21 out of 32 apes chose correctly), $\chi^2(1, N = 32) = 3.125, p = .10, C = .30$, it was apparent by the fourth trial as 26 of the 32 apes chose the top container, $\chi^2(1, N = 32) = 12.500, p < .001, C = .53$. Seventy-five percent of the apes selected the correct cup in at least four of the first five trials, $\chi^2(1, N = 32) = 8.000, p < .01, C = .45$, whereas 97% reached this criterion within 10 trials, $\chi^2(1, N = 32) = 28.125, p < .001, C = .68$.

The four ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, N = 32) = 1.440, p = .71$, or their preference in the first, $\chi^2(1, N = 32) = 1.588, p = .72, C = .31$, and fourth trials, $\chi^2(1, N = 32) = 5.505, p = .14$,

Table 2
Percentage of Correct Responses for the First, Fourth, and All Trials in Each of the Release Conditions of Experiment 1

	Far			Near		
	1 st trial	4 th trial	All	1 st trial	4 th trial	All
Orangutan ($n = 11$)	45.5	63.6	63.8	54.5	81.8	72.8
Chimpanzee ^a ($n = 11$)	63.6	72.7	70.5	63.6	100	86.3
Gorilla ($n = 10$)	70.0	80.0	75.0	80.0	60.0	77.5
Total	59.4	71.9	69.6	65.6	81.3	78.9

^a includes chimpanzees and bonobos.

$C = .54$. The same pattern appeared when comparing only five individuals of each species. The four ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, n = 20) = 5.523, p = .14$, or in their preference in the first, $\chi^2(3, n = 20) = 5.000, p = .32, C = .63$, or the fourth trial, $\chi^2(3, n = 20) = 5.000, p = .25, C = .63$. Apes reliably differed in their overall preference for the top container across age, Spearman correlation $r(32) = -.35, p < .05$. A closer look at the data revealed that the age difference did not follow a developmental trend: the oldest apes (over 25 years) made most errors, followed by apes of middle age (from 5 to 12 years). The fewest errors were made by the youngest age group (4 years) and by apes aged from 13 to 24 years.

Comparison Between Vertical Conditions

Although all apes performed better in the near drop condition, there were no overall significant differences between the far and near conditions in the percentage of correct choices, Wilcoxon's test: $z = 1.327, p = .19$. Similarly, there were no significant differences between conditions in the first, $\chi^2(1, N = 32) = 1.347, p = .28, C = .20$, or fourth trial, $\chi^2(1, N = 32) = 1.748, p = .31, C = .23$. Although more apes reached the criterion of 4 out of 5 correct trials faster in the near condition compared to the far condition, this difference was not significant, McNemar test: $p = .5$.

Discussion

Overall, great apes exhibited no gravity bias, even if only the first trial was considered. Instead, apes showed a significant preference for the upper location in the fourth trial. Even if a weak tendency to choose the lower location initially existed, this tendency cannot be characterized as inflexible and immune to learning. Indeed, most apes consistently selected the upper location within 10 trials. Orangutans appeared to be more error prone than African apes in the first trials of first task exposure.

Contrary to our expectations, varying the length of the trajectory of the falling object had little effect on the selection bias. Even if, as a tendency, all species more often selected the lower location, this tendency was not statistically reliable, except that apes reached criterion faster in the near condition. Overall, there were few age-related differences. Only in the near condition we found an age difference in the overall preference for the top container. However, this difference was unsystematic and did not reveal a clear developmental trend, since the oldest apes made the most and the youngest apes the fewest errors. Rather, this difference might reflect individual or motivational aspects.

The initial performance of the apes supported none of the hypotheses postulated (since they showed no cup preference). Both, gravity and shelter hypotheses were not supported since they predicted a bottom location bias. Also, proximity and solidity hypotheses were not clearly supported. Even though the apes preferred the upper location from the fourth trial on and most of them learned to select the upper location in the course of the task, it remains doubtful if this can be interpreted as evidence for the proximity or solidity hypotheses, which both predicted an upper location search from the first trial on. It is possible that a tendency to search in the upper cup is present in some apes, but that this tendency is masked in a vertical version of the task because

vertical displacements also evoke gravity related errors. Recall that the gravity error is thought to be overcome because the ability to suppress the error increases during development. In this case, the selection of a cup results out of two processes: first, the intention to search a specific location, and second, the ability to suppress a perseverating tendency to reach for another location. It is possible that apes represent an intermediate developmental state where the inhibitory capacities are considerably but not sufficiently heightened, and as a consequence neither the error nor the correct choice is consistently traceable. If this is true, then the tendency to select the cup specified by solidity/proximity constraints should become apparent, once factors affecting the ability to suppress gravity related responses (i.e., the vertical displacement) are removed. The following experiment confronts the apes with the same tasks but presented along the horizontal dimension. If apes' tendency to select the cup specified by solidity or proximity was masked due to inhibitory failure they should consistently select this cup when vertical displacement is removed.

Experiment 2: Length of the Horizontal Trajectory

This experiment investigated whether apes showed a preference for the near location compared to the far location when the task was presented on the horizontal plane. Such a preference would be consistent with both the proximity and the solidity hypotheses. The shelter hypothesis makes no specific prediction. Participants witnessed either a long or a short horizontal trajectory before the object moved behind the screen. This manipulation allowed us to assess whether the length of the trajectory had any effect on the cup selection. The gravity hypothesis predicts that apes should perform better on horizontal than vertical trials (because gravity errors are restricted to vertical object displacements) and that—in contrast to vertical trials—the length of object trajectory should not influence search behavior on horizontal trials (if the ability to suppress gravity errors is affected by seeing a long vertical fall).

Procedure

Figure 2 depicts the cup arrangement for this experiment. One cup was placed at the far end of the surface, with the other cup 19 cm away from it. The cups' openings faced the end of the surface where the object was released. All horizontal displacements were presented in two spatial arrangements: Half of the participants had the cup openings facing to the right side of the surface (hence the experimenter rolled the grape from the right side) and half to the left side. All participants were run on two baseline and two test conditions, followed by a control condition.

Before receiving the horizontal tests, participants first had to pass two baseline conditions, whose rationale was identical to those in Experiment 1.

Horizontal baseline condition. A grape was placed randomly in either cup while the ape watched. If the ape selected the baited cup in nine out of 10 trials it moved on to the *horizontal switch condition*.

Horizontal switch condition. The grape was first placed in the same cup on five consecutive trials, in the sixth trial the grape was put in the other cup, followed by four trials in which it was randomly placed. If the ape correctly switched to the new location in the sixth trial, it moved on to the test conditions.

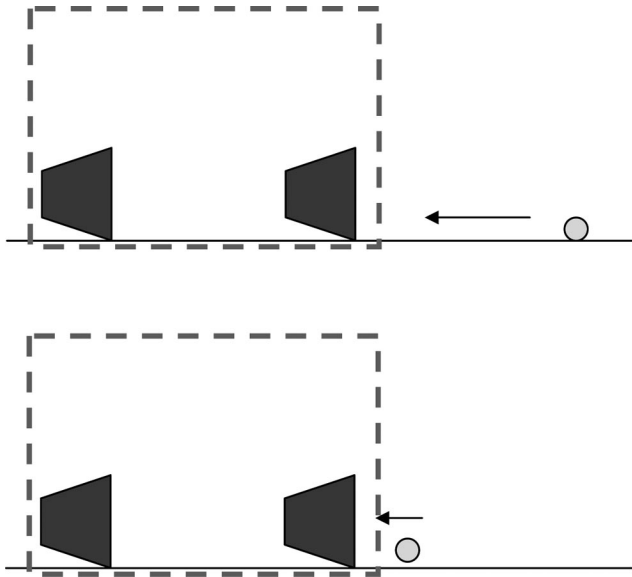


Figure 2. Set up for the test conditions in Experiment 2. The horizontal far release condition is depicted in the upper section, the horizontal near release condition on the lower section. The green dotted line indicates the screen.

Upon passing both baseline conditions (something that all apes did), they received the following two test conditions in counter-balanced order (see Figure 2) and a final control condition:

Horizontal far release condition. The cups were arranged along the horizontal plane as in the baseline condition, but a screen was raised and a grape held out to the ape. If the ape tracked the grape, the experimenter placed the grape on the platform opposite to the cups and rolled it behind the screen (25 cm from the beginning of the screen). Then the screen was removed and the platform was pushed forward toward the mesh. If the subject selected the correct front cup, the experimenter gave her the grape. If the ape chose the wrong location, the experimenter showed the empty cup to the ape. Then the cup was placed in its original position and the ape was allowed to make a new choice. This was repeated until the ape found the grape or stopped searching.

Horizontal near release condition. This condition was identical to the *horizontal far release condition* except that the grape was released (start point of rolling) directly in front of the screen.

Horizontal far release control. This condition was identical to the *horizontal far release condition* except that subjects received the opposite spatial arrangement to the one they had previously experienced. This allowed us to confirm that no side bias was responsible for the observed behavior.

Results

First Exposure to the Task (Including Either Far or Near Release Conditions)

Again, we analyzed apes' choice in the first trial of their first exposure to the task (either far or near drop condition) to check for potential biases that might be washed out in the course of the task. Apes preferentially selected the correct container in the first trial, $\chi^2(1, N = 32) = 21.125, p < .001, C = .63$, with 29 of 32 apes choosing correctly. The same is true for the 4th trial, $\chi^2(1, N = 32) = 18.00, p < .001, C = .60$, with 28 of 32 choosing correctly. Ninety-four percent of the apes selected the correct cup in at least four of the first five trials, $\chi^2(1, N = 32) = 24.500, p < .001, C = .66$, whereas all of them reached this criterion within 10 trials, $p < .001$.

The four ape species did not differ in their overall cup preference, Kruskal-Wallis test: $\chi^2(3, N = 32) = 2.550, p = .46$, or in their preference in the first, $\chi^2(3, N = 32) = 2.468, p = .53, C = .38$, and the fourth trial, $\chi^2(3, N = 32) = 3.269, p = .43, C = .43$. The same was found when comparing only five individuals of each species. The four ape species did not differ in their overall cup preference, Kruskal-Wallis test: $\chi^2(3, n = 20) = 2.766, p = .44$, or in the first, $\chi^2(3, n = 20) = 3.158, p = 1.0, C = .52$, and the fourth trial, $\chi^2(3, n = 20) = 2.222, p = 1.0, C = .45$. Correlating age with percentage of correct choices revealed that apes did not differ in their overall preference for the front container across age, Spearman correlation $r(32) = -.06, p = .73$.

Horizontal Far Release Condition

Table 3 presents the percentage of subjects that selected the correct container in the first and fourth trials (far and near conditions) for each species. Overall, apes selected the near container in 92% of the trials, Wilcoxon's test: $z = 5.077, p < .001$; see Table 3. This preference for the near container was observed in both the first, $\chi^2(1, N = 32) = 28.125, p < .001, C = .68$, and fourth trial, $\chi^2(1, N = 32) = 15.125, p < .001, C = .57$. Ninety-four percent of the apes selected the correct cup in at least four of the first five

Table 3
Percentage of Correct Responses for the First, Fourth, and All Trials in Each of the Release Conditions of Experiment 2

	Far			Near		
	1 st trial	4 th trial	All	1 st trial	4 th trial	All
Orangutan ($n = 11$)	90.9	81.8	86.3	90.9	90.9	91.0
Chimpanzee ^a ($n = 11$)	100	81.8	95.5	100	100	100
Gorilla ($n = 10$)	100	90.0	92.5	70.0	90.0	90.0
Total	96.9	84.4	91.4	87.5	93.8	93.8

^a includes chimpanzees and bonobos.

trials, $\chi^2(1, N = 32) = 24.500, p < .001, C = .66$, whereas all of them reached this criterion within 10 trials, $p < .001$.

The four ape species did not differ in their overall preference for the near container, Kruskal-Wallis test: $\chi^2(3, N = 32) = 2.551, p = .51$, or in their preference in the first, $\chi^2(3, N = 32) = 1.909, p = 1.00, C = .34$, or the fourth trial, $\chi^2(3, N = 32) = 2.565, p = .58, C = .39$. The same pattern appeared when comparing only five individuals of each species. The four ape species did not differ in their overall preference for the front container, Kruskal-Wallis test: $\chi^2(3, n = 20) = 4.670, p = .20$, or in their preference in the first, $\chi^2(3, n = 20) = 3.158, p = 1.00, C = .52$, or the fourth trial, $\chi^2(3, n = 20) = 1.176, p = 1.00, C = .33$. Age groups did not differ in their overall preference for the front container across age groups, Spearman correlation $r(32) = -.19, p = .29$.

Horizontal Near Release Condition

Overall, apes selected the near container in 94% of the trials, Wilcoxon's test: $z = 5.20, p < .001$; see Table 3. This preference for the near container was observed in both the first, $\chi^2(1, N = 32) = 18.000, p < .001, C = .60$, and fourth trial, $\chi^2(1, N = 32) = 24.500, p < .001, C = .66$, with 28 and 30 of 32 apes searching the correct location. Ninety-seven percent of the apes selected the correct cup in at least four of the first five trials, $\chi^2(1, N = 32) = 28.125, p < .001, C = .68$, whereas all of them reached this criterion within 10 trials, $p < .001$.

Ape species did not differ in their overall preference for the near container, Kruskal-Wallis test: $\chi^2(3, N = 32) = 4.828, p = .19$, or in their preference in the first, $\chi^2(3, N = 32) = 4.348, p = .28, C = .49$, or the fourth trial, $\chi^2(3, N = 32) = 1.090, p = 1.00, C = .26$. The same pattern appeared when comparing only five individuals of each species. The four ape species did not differ in their overall preference for the near container, Kruskal-Wallis test: $\chi^2(3, n = 20) = 7.300, p = .10$, or in their preference in the first, $\chi^2(3, n = 20) = 3.158, p = 1.00, C = .52$, or the fourth trial, $\chi^2(3, n = 20) = 2.222, p = 1.00, C = .45$. The overall preference for the near container did not differ across age, Spearman correlation $r(32) = -.12, p = .50$.

Comparison Between Horizontal Conditions

There were no overall significant differences between the far and near conditions in the percentage of correct choices, Wilcoxon's test: $z = .758, p = .47$. Similarly, there were no significant differences between conditions in the first, $\chi^2(1, N = 32) = 0.147, p = 1.00, C = .07$, or fourth trial, $\chi^2(1, N = 32) = 0.395, p = 1.00, C = .11$. Finally, there were no side preferences: apes performed correctly in the first trial of opposite side control condition, $\chi^2(1, N = 32) = 18.000, p < .001, C = .60$.

Discussion

All ape species reliably selected the front cup in the first trial and continued to do so in subsequent trials. So, once the vertical displacement was removed, apes consistently selected the correct cup. This suggests that the errors observed in the vertical version of Experiment 1 were due to the failure of inhibiting gravity responses. These results support the gravity hypothesis, which predicted better performance in horizontal than vertical conditions.

The gravity hypothesis also predicted no effects for the length of the reward's visible trajectory before it disappeared behind the screen, which was the case. In contrast to vertical displacements apes reached criterion equally fast in both conditions. However, also in vertical displacements the length of the trajectory had little impact overall.

Also, both the proximity and the solidity hypotheses were supported because they predicted a front cup selection. However, it seems that the tendency to select the cup specified by solidity/proximity is in conflict with gravity related responses and, hence, only on horizontal tasks it guides performance reliably from the first trial on.

Experiment 3: Degree of Misalignment in the Vertical Displacement

Experiment 3 returned to vertical displacements by investigating the effect of misalignment between the cups, that is, the bottom cup was not placed straight under the dropping point, but shifted to the side. If our conclusion that apes' errors were associated with gravity in Experiment 1 was correct, then a cup deviation should lead to more correct responses.

In one condition, subjects experienced a slight misalignment between the cups while in another they experienced a substantial misalignment. Recall that the gravity hypothesis predicts that the larger the deviation of the bottom cup from the object's falling trajectory, the better should be the subjects' search performance. All other hypotheses predict no effect of cup misalignment on cup selection.

Procedure

Prior to the test, apes who were first tested with Experiment 3 received two baseline conditions that served to a) familiarize participants with the apparatus, b) exclude baseline preferences for specific cup locations, and c) exclude inhibitory problems influencing the search behavior (the vertical baseline conditions are presented in Experiment 1, see page 13 for a detailed description of the procedure).

Subjects received two test conditions (see Figure 3) in counter-balanced order and one control condition.

Vertical slight misalignment condition. The setup and procedure was identical to the far drop condition of Experiment 1 except that the bottom cup was misaligned by 3 cm.

Vertical large misalignment condition. This condition was identical to the slight misalignment condition except that both cups were moved away from the table center and misaligned by 19 cm from each other. Half of the participants were presented with a top-left, down-right configuration, half of them with the opposite top-right, down-left configuration.

Vertical misalignment control. We retested all apes on the large misalignment condition but with the opposite cup configuration to check for side biases within participants.

Results

First Exposure to the Task (Including Either Slight or Large Misalignment Conditions)

Apes preferentially selected the correct container in the first trial, $\chi^2(1, N = 32) = 8.000, p < .01, C = .45$, with 24 of 32 apes

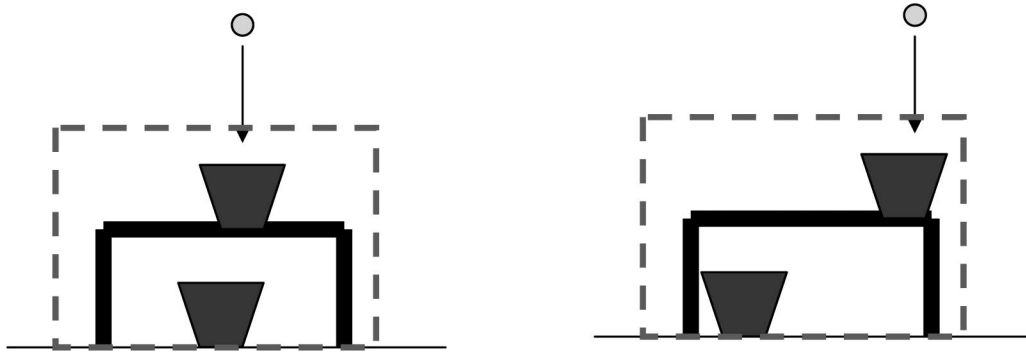


Figure 3. Set up for the test conditions in Experiment 3. The vertical slight misalignment is depicted on the left side, the vertical great misalignment condition is depicted on the right side. The green dotted line indicates the screen.

choosing correctly. The same is true for the 4th trial, $\chi^2(1, N = 32) = 12.500, p < .001, C = .53$, with 26 of 32 choosing correctly. Seventy-five percent of the apes selected the correct cup in at least four of the first five trials, $\chi^2(1, N = 32) = 8.000, p < .01, C = .45$, whereas 90% reached this criterion within 10 trials, $\chi^2(1, N = 32) = 21.125, p < .001, C = .63$.

Ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, N = 32) = 1.92, p = .61$, or in their preference in the first, $\chi^2(3, N = 32) = 4.760, p = .20, C = .51$, or the fourth trial, $\chi^2(3, N = 32) = 3.772, p = .35, C = .46$. The same pattern appeared when comparing only five individuals of each species. The four ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, n = 20) = 7.308, p = .05$, or in their preference in the first, $\chi^2(3, n = 20) = 7.692, p = .09, C = .75$, or the fourth trial, $\chi^2(3, n = 20) = 2.933, p = .68, C = .50$. Correlating age with percentage of correct choices revealed that the overall preference for the top container did not differ across age, Spearman correlation $r(32) = -.13, p = .47$.

Vertical Slight Misalignment Condition

Overall, apes selected the top container in 72% of the trials, Wilcoxon's test: $z = 2.842, p < .01$; see Table 4). This preference for the top container was observed in both the first, $\chi^2(1, N = 32) = 6.125, p < .05, C = .40$, and fourth trial, $\chi^2(1, N = 32) = 10.125, p < .01, C = .49$. Seventy-five percent of the apes selected the correct cup in at least four of the first five trials, $\chi^2(1, N =$

$32) = 8.000, p < .01, C = .45$, whereas 88% reached this criterion within 10 trials, $\chi^2(1, N = 32) = 18.000, p < .001, C = .60$.

Ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, N = 32) = 1.114, p = .79$, or in their preference in the first, $\chi^2(3, N = 32) = 2.348, p = .51, C = .37$, or the fourth trial, $\chi^2(3, N = 32) = 4.667, p = .24, C = .50$. The same pattern appeared when comparing only five individuals of each species. The four ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, n = 20) = 5.331, p = .15$, or in their preference in the first, $\chi^2(3, n = 20) = 5.000, p = .32, C = .63$, or the fourth trial, $\chi^2(3, n = 20) = 4.762, p = .38, C = .62$. The overall preference for the top container did not differ across age, Spearman correlation $r(32) = -.11, p = .55$.

Vertical Large Misalignment Condition

Overall, apes selected the top container in 87% of the trials, Wilcoxon's test: $z = 4.588, p < .001$; see Table 4. This preference for the top container was observed in both the first, $\chi^2(1, N = 32) = 10.125, p < .01, C = .49$, and fourth trial, $\chi^2(1, N = 32) = 24.500, p < .001, C = .66$. Eighty-four percent of the apes selected the correct cup in at least four of the first five trials, $\chi^2(1, N = 32) = 15.125, p < .01, C = .57$, whereas 94% reached this criterion within 10 trials, $\chi^2(1, N = 32) = 24.500, p < .001, C = .66$.

Ape species did not differ in their overall preference for the top container, Kruskal-Wallis test: $\chi^2(3, N = 32) = 0.895, p = .83$, or in their preference in the first, $\chi^2(3, N = 32) = 2.211, p = .59$,

Table 4
Percentage of Correct Responses for the First, Fourth, and All Trials in Each of the Misalignment Conditions of Experiment 3

	Slight			Large		
	1 st trial	4 th trial	All	1 st trial	4 th trial	All
Orangutan ($n = 11$)	63.6	63.6	61.3	63.6	90.9	79.5
Chimpanzee ^a ($n = 11$)	63.6	81.8	72.8	81.8	90.9	88.8
Gorilla ($n = 10$)	90.0	90.0	82.5	90.0	100	92.5
Total	71.9	78.1	71.9	78.1	93.8	86.8

^a includes chimpanzees and bonobos.

$C = .36$, or the fourth trial, $\chi^2(3, N = 32) = 2.743, p = .53, C = .40$. The same pattern appeared when comparing only five individuals of each species. The four ape species did not differ in their overall preference for the upper cup, Kruskal-Wallis test: $\chi^2(3, n = 20) = 4.653, p = .21$, or in their preference in the first, $\chi^2(3, n = 20) = 5.934, p = .20, C = .68$, or the fourth trial, $\chi^2(3, n = 20) = 2.222, p = 1.00, C = .45$. The overall preference for the top container did not differ across age, Spearman correlation $r(32) = -.01, p = .95$.

Comparison Between Vertical Misalignment Conditions

Overall, apes performed better in the large misalignment condition compared to the slight misalignment condition, Wilcoxon's test: $z = 2.128, p < .05$. This difference, however, was not statistically significant if the first, $\chi^2(1, N = 32) = 3.732, p = .08, C = .32$, or fourth trials, $\chi^2(1, N = 32) = 0.987, p = .40, C = .17$, were considered separately.

Comparison Between Horizontal and Vertical Tasks

Apes performed significantly better in the horizontal condition (92.6% correct) compared to the vertical condition of Experiments 1, Wilcoxon's test: $z = 3.485, p < .01, 74.6\%$ correct, and Experiment 3, Wilcoxon's test: $z = 2.999, p < .01, 79.3\%$ correct. The difference between the horizontal task and the vertical tasks was evident in the first test trial, Wilcoxon's test: $z = 2.67, p < .05$ in both cases. In contrast, there were no significant differences between the vertical conditions in Experiments 1 and 3 both overall, Wilcoxon's test: $z = .851, p = .40$, and in the first trial, Wilcoxon's test: $z = 1.667, p = .18$.

Influence of Test Order in Vertical Experiments 1 and 3

Order of task administration had no statistically significant effect on the overall performance in Experiment 1, Mann-Whitney test: $z = 1.94, p = .06$, and Experiment 3, Mann-Whitney test: $z = .099, p = .93$. Breaking down the tasks into their conditions and analyzing the first trial also revealed no significant effect of order of presentation: Fisher tests: far: $\chi^2(1, N = 32) = 1.499, p = .29, C = .21$; near: $\chi^2(1, N = 32) = 0.372, p = .71, C = .10$; slight misalignment: $\chi^2(1, N = 32) = 0.002, p = 1.0, C = .01$; and large misalignment: $\chi^2(1, N = 32) = 0.003, p = 1.0, C = .01$. Similarly, there was no effect on those conditions in the fourth trial: Fisher tests: far: $\chi^2(1, N = 32) = 0.552, p = .69, C = .13$; slight misalignment: $\chi^2(1, N = 32) = 0.839, p = .43, C = .16$; large misalignment: $\chi^2(1, N = 32) = 0.034, p = 1.0, C = .03$; except for the near condition (of Experiment 1) in which subjects performed significantly better after they had received Experiment 3, Fisher test: near: $\chi^2(1, N = 32) = 5.744, p < .05, C = .39$.

Discussion

Apes searched correctly (in the top location) in both conditions from the first trial regardless of the degree of misalignment between the bottom cup and the release point. Although apes' overall performance was better in the large misalignment compared to the slight misalignment condition, this difference was not apparent in the first trial.

Searching the top location is consistent with both the proximity and solidity hypotheses and squarely against the shelter hypothesis that predicted search under the table-shelf in both experimental conditions. Top location search does not support the gravity hypothesis either. However, the gravity hypothesis is the only one that predicted the observed performance improvement due to an increase in cup misalignment.

General Discussion

Initially, apes showed neither a preference for the upper location nor a robust gravity bias. In this respect, they clearly differ from rhesus monkeys, which showed a gravity-bias in analogous tasks (Hauser, 2001; Southgate & Gómez, 2006). Rather, they display a pattern similar to that of 2.5-year-old children who failed in this task even though they were not particularly biased to search in the incorrect down location. Orangutans appeared to be more error prone than African apes in the first trials of initial task exposure. It is unclear why this was the case. Perhaps the more arboreal Asian apes might have experienced greater difficulties suppressing gravity-driven responses than African apes. However, even if a tendency to gravity biased cup selection initially existed, the apes quickly learned to inhibit it, a persistent prepotent gravity bias was not observed. These results are in line with our findings with the tubes task (Cacchione & Call, 2008) and contrast those of Tomonaga et al. (2007) who found a gravity bias in a prediction task. However, gravity related factors affected the probability for a gravity response. While apes initially failed to select the correct cup in a vertically straight arrangement, they reliably selected the cup specified by solidity/proximity in horizontal displacements. Also, apes reliably select the correct cup when tested with a deviated cup arrangement and were less error-prone when the misalignment between search locations was increased. This is clearly in favor of the gravity hypothesis. Again, these findings are in line with our results with the tubes task, where a reappearance of gravity biased search was also observed under circumstances affecting inhibition.

Taken together, these findings suggest that apes were influenced by both a tendency to select the upper cup and a tendency to react to gravity constraints when searching for an invisibly displaced object. It is unclear, however, what led to correct cup selection in horizontal and misaligned arrangements. It is possible that the apes can infer that solidity prevents an object from moving through the table and that, as a consequence, the object can be found in the upper cup. However, even though a proximity strategy cannot account for the influence of gravity associated factors, the present results cannot rule out that proximity rather than solidity could be a valid explanation for searching the upper location. Future studies will be needed to disentangle these two hypotheses. Compared to the other hypotheses, the shelter hypothesis received very little empirical support.

At a first glance, it seems difficult to understand why apes do not show robust evidence of a gravity error, yet perform worse on versions of the task that would accentuate a gravity bias. However, recall that development gives rise to both, the capacity to infer where the target object is hidden and the capacity to search where it is expected by inhibiting unwarranted responses. Apes might have acquired sufficient inhibitory control not to show a reliable bias and yet gravity related factors (i.e., vertical displacements,

straight cup arrangement) may still impair their inhibitory capacities. Given the large developmental changes that have been described in humans for this task, it may seem surprising not to find any developmentally relevant age effects in the present study. However, our sample might not be adequate to trace developmental changes out of two reasons. First and foremost, our sample lacked apes younger than 4 years of age. We know from human infants that the main developmental changes occur before 3 years of age and apes were found to develop cognitive capacities in even greater speed than human infants (Gomez, 2005). Second, our sample is not evenly distributed across ages. Further studies should address these problems.

Recently, Santos (2004) found evidence for a two-systems account of initial knowledge originally proposed by Scholl and Leslie (1999), which might explain why apes never gain as much inhibitory control as infants in tasks that require reasoning about mechanical interactions. The idea is that initial knowledge consists of a spatiotemporal system (which serves to track objects over space and time) and a contact-mechanical system (which serves to reason about mechanical interactions between objects). Note that all events presented in our study involved mechanical interactions between objects. Santos (2004) observed that rhesus macaques performed well in action tasks involving spatiotemporal but failed in similar tasks involving contact-mechanical knowledge. In contrast to the automatic and modular operation of spatiotemporal knowledge, the contact-mechanical reasoning is less encapsulated and takes in outside information during processing. As a result, contact-mechanical reasoning seems more susceptible to outside influence in the form of perseverative biases, naïve theories, or problems with task demands. Since all our action tasks requested mechanical reasoning, this might explain why in some instances gravity related factors disrupted the apes' generally good performance. Even if apes are apparently more apt to cope with mechanical interactions than monkeys, they might still face more problems in this respect than three-year-old human infants.

One potential caveat of our task is that subjects received a baseline in which they retrieved several times the reward from both the upper and lower cups the same number of occasions, thus making our data not comparable to previous studies. However, if simply letting subjects retrieve the reward from both locations equally often is enough to overcome a prepotent gravity bias, we have to say that such a bias, if existed, was not very strong in the first place. Certainly, it was not as strong as that displayed by rhesus macaques (Hauser, 2001; Southgate & Gómez, 2006).

From a comparative point of view, it is interesting that apes appear to have overcome a robust bias while rhesus macaques display a strong gravity bias (Hauser, 2001). While children are supposed to grow out of the gravity bias because they acquire the capacity to inhibit prepotent responses, monkeys apparently never acquire sufficient inhibitory skills. It is tempting to suggest that apes have evolved the cognitive mechanisms that allow them—at least to some extent—to overcome the strong gravity bias displayed by monkeys, which would represent the ancestral state. Drawing such conclusion, however, would be premature because of two reasons. First, we know nothing about the behavior of other monkey species in the table task. One could argue that great apes represent the ancestral state and rhesus macaques have recently evolved the gravity bias, perhaps in response to increased feeding competition (Southgate & Gómez, 2006). Consequently, future

research should be devoted to investigate whether other monkey species also display a gravity-biased search. Gibbons, which are classified between monkeys and great apes, would be particularly interesting species. Hoolock gibbons, in particular, have recently passed a task in which they have to pull in a reward with a tool while avoiding a trap in which the reward may fall (Cunningham, Anderson, & Mootnick, 2006). Second, Tomonaga et al. (2007) found a persistent gravity bias in chimpanzees, which complicates the view that apes (in contrast to monkeys) evolved the capacity to inhibit prepotent responses. It is possible, however, that the need to predict the landing location of an object before it is dropped is more likely to trigger a gravity response than to search for the object after dropping. Future research should address the question why and under what circumstances gravity biased search can be observed in a specific task and why and under what circumstances an already overcome gravity bias may eventually reappear.

References

- Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 47–83). Oxford: Blackwell.
- Baillargeon, R. (2004). Infants reasoning about hidden objects: Evidence for event-general and event-specific expectations. *Developmental Science*, 7, 391–424.
- Berthier, N., DeBlois, S., Poirer, C. R., Novak, M. A., & Clifton, R. K. (2000). Where's the ball? Two- and three-year-olds reason about unseen events. *Developmental Psychology*, 36, 394–401.
- Bliss, J., Ogborn, J., & Whitelock, D. (1989). Secondary school pupils' common sense theories of motion. *International Journal of Science Education*, 11, 261–272.
- Butler, S. C., Berthier, N., & Clifton, R. K. (2002). Two-year-old's search strategies and visual tracking in a hidden displacement task. *Developmental Psychology*, 38, 581–590.
- Cacchione, T., & Call, J. (2008). *Do great apes refer to the tube as a causal device? Intuitions about gravity and solidity in four great ape species*. Manuscript in preparation.
- Cacchione, T., & Krist, H. (2004). Recognizing impossible object relations: Intuitions about support in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 118, 140–148.
- Cashon, C. H., & Cohen, L. B. (2000). Eight month-old infants' perception of possible and impossible events. *Infancy*, 1, 429–446.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Erlbaum.
- Cunningham, C. L., Anderson, J. R., & Mootnick, A. R. (2006). Object manipulation to obtain a food reward in hoolock gibbons, *Bunopithecus hoolock*. *Animal Behaviour*, 71, 621–629.
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, 10, 483–527.
- Gómez, J. C. (2005). Species comparative studies and cognitive development. *Trends in Cognitive Sciences*, 9, 118–125.
- Hauser, M. D. (2001). Searching for food in the wild. A nonhuman primate's expectations about invisible displacement. *Developmental Science*, 4, 84–93.
- Hauser, M. D. (2003). Knowing about knowing. Associations between perception and action systems over evolution and during development. *Annual New York Academy of Sciences*, 1, 1–25.
- Hauser, M. D., Williams, T., Kralik, J. D., & Moskowitz, D. (2001). What guides a search for food that has disappeared? Experiments on cotton-top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology*, 115, 140–151.
- Hood, B. M. (1995). Gravity rules for 2–4-year-olds? *Cognitive Development*, 10, 577–598.

- Hood, B. M. (1998). Gravity does rule for falling events. *Developmental Science, 1*, 59–64.
- Hood, B. M., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events. *Child Development, 71*, 1540–1554.
- Hood, B. M., Cole-Davies, V., & Dias, M. (2003). Looking and search measures of object knowledge in pre-school children. *Developmental Psychology, 39*, 61–70.
- Hood, B. M., Hauser, M. D., Anderson, L., & Santos, L. (1999). Gravity biases in a nonhuman primate? *Developmental Science, 2*, 35–41.
- Hood, B. M., Santos, L., & Fieselman, S. (2000). Two year-old's naïve predictions for horizontal trajectories. *Developmental Science, 3*, 328–332.
- Hood, B. M., Wilson, A., & Dyson, S. (2006). The effect of divided attention on inhibiting the gravity error. *Developmental Science, 9*, 303–308.
- Kaiser, M. K., Proffitt, D. R., & McCloskey, M. (1985b). The development of beliefs about falling objects. *Perception & Psychophysics, 38*, 533–539.
- Karin-D'Arcy, M. R., & Povinelli, D. (2002). Do chimpanzees know what each other see? A closer look. *International Journal of Comparative Psychology, 15*, 21–54.
- Kim, I. K., & Spelke, E. S. (1992). Infants' sensitivity to effects of gravity on visible object motion. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 385–393.
- Kim, I. K., & Spelke, E. S. (1999). Perception and understanding of effects of gravity and inertia on object motion. *Developmental Science, 2*, 339–362.
- Krist, H. (2000). Development of naïve beliefs about moving objects: The straight-down belief in action. *Cognitive Development, 15*, 281–308.
- Mash, C., Keen, R., & Berthier, N. E. (2003). Visual access and attention in two-year-old's event reasoning and object search. *Infancy, 4*, 371–388.
- Osthaus, B., Slater, A. M., & Lea, S. E. G. (2003). Can dogs defy gravity? A comparison with the human infant and a non-human primate. *Developmental Science, 6*, 489–497.
- Santos, L. R., & Hauser, M. D. (2002). A non-human primates understanding of solidity: Dissociation between seeing and acting. *Developmental Science, 5*, F1–F7.
- Santos, L. R. (2004). 'Core Knowledges': a dissociation between spatio-temporal knowledge and contact-mechanisms in a non-human primate? *Developmental Science, 7*, 167–174.
- Santos, L. R., Seelig, D., & Hauser, M. D. (2006). Cotton-top tamarins' (*Saguinus oedipus*) expectations about occluded objects: A dissociation between looking and reaching tasks. *Infancy, 9*, 147–171.
- Scholl, B. J., & Leslie, A. M. (1999). Explaining the infant's object concept beyond the perception/cognition dichotomy. In E. Lepore & Z. Pylyshyn (Eds.), *What is cognitive science?* (pp. 26–73). Oxford: Blackwell.
- Southgate, V. (2005). A comparative study of the understanding of invisible object displacement in macaque monkeys (*Macaca mulatta* and *arctoides*) and children (*Homo sapiens*). Unpublished doctoral dissertation, University of St. Andrews.
- Southgate, V., & Gómez, J. C. (2006). Searching beneath the shelf in macaque monkeys: Evidence for a gravity bias or a foraging bias? *Journal of Comparative Psychology, 120*, 314–321.
- Southgate, V., Gómez, J. C., Fox, S., & Meints, K. (2005). *Eye gaze reveals two-year-olds' ability to predict outcomes*. Paper presented at the 3rd International Workshop for Young Psychologists on Evolution and Development of Cognition, Kyoto, Japan.
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*(4), 605–632.
- Tomonaga, M., Imura, T., Mizuno, Y., & Tanaka, M. (2007). Gravity bias in young and adult chimpanzees (*Pan troglodytes*): Tests with a modified opaque-tubes task. *Developmental Science, 10*, 411–421.

Received January 18, 2008

Revision received July 14, 2008

Accepted July 22, 2008 ■