

## RESEARCH ARTICLE

## Great Apes Use Weight as a Cue to Find Hidden Food

CORNELIA SCHRAUF<sup>1\*</sup> AND JOSEP CALL<sup>2</sup><sup>1</sup>Department of Behavioural Biology, University of Vienna, Vienna, Austria<sup>2</sup>Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Bonobos (*Pan paniscus*;  $n = 5$ ), orangutans (*Pongo pygmaeus abelii*;  $n = 6$ ), and a gorilla (*Gorilla gorilla gorilla*;  $n = 1$ ) were presented with two opaque cups, one empty and one baited (containing two bananas). Subjects had to independently gain weight information about the contents of the cups to find the hidden food. Six apes attained above chance level within a total of 16 trials. Successful subjects spontaneously adopted the method of successively lifting the cups and thus comparing their weight before making a choice. Prior to testing, these apes had participated in a weight discrimination task. To rule out that a subject's good performance was influenced by previous experience in weight experiments, we ran a second test in which the same task was presented to a group of chimpanzees (*Pan troglodytes*;  $n = 9$ ) who were naïve to weight experiments. These subjects also participated in an additional test condition in which the same problem was presented based on learning to associate arbitrary visual stimuli. The results show that experience did not affect performance because the nine naïve subjects were equally able to find the food when the task stimuli held a causal relation (i.e. weight indicates the hidden food). Interestingly, only one of the naïve subjects solved the task when the task elements held an arbitrary relation (i.e. certain visual pattern indicates food). Our results confirm previous findings that apes perform better in problems grounded on causal compared to arbitrary relations. *Am. J. Primatol.* 73:323–334, 2011. © 2010 Wiley-Liss, Inc.

**Key words:** weight discrimination; kinesthetic perception; causality

## INTRODUCTION

Contemporary psychology often assumes that animals can learn to associate stimuli and responses, but possess little knowledge about the relation between observed events [e.g. Povinelli, 2000; Thorndike, 1898; Woodward, 2007]. Recent studies, however, show that the simple formation of arbitrary connections between certain events and responses is insufficient to explain various phenomena [e.g. Blaisdell et al., 2006; Call, 2004; Seed et al., 2006]. Seed and Call [2009] argued that certain species such as apes and corvids perceive functional relations between elements, and do not merely register their co-occurrence. Furthermore, Call [2006a] noted that establishing stimuli–response associations is much harder when the elements of a task hold an arbitrary relation rather than a causal/functional relation. More specifically, the relation between a cue and a reward can be processed in two different ways, depending on whether the cue co-varies with the presence of food or the cue is caused by the presence of food. One important question is whether animals are sensitive to this distinction by responding differently to cues that are based on co-variation as opposed to causation.

One context in which this discrepancy has been repeatedly observed is when subjects have to use information from an event to infer the location of

food [e.g. Call, 2004, 2006a,b; Hanus & Call, 2008]. For instance, Call [2004] presented great apes with two opaque containers and gave visual and auditory information about the content of the cups, only one of which contained food. When the cue and the food held a causal relation, i.e. shaking the cup produced either a noise or silence depending on whether there was food inside or not, subjects effectively used the auditory cue to find the food. In contrast, when the cue and the food held an arbitrary relation, i.e. tapping the cup produced the noise, subjects failed to use the cue to find the food. These results are remarkable given that both cues had the same predictive power because they always reliably indicated the presence of food.

A sensitivity to the causal–arbitrary distinction may not be restricted to apes. Sabbatini and

Contract grant sponsor: Austrian Academy of Science to Cornelia Schrauf.

\*Correspondence to: Cornelia Schrauf, Department of Behavioural Biology, University of Vienna, Althanstrasse 14, A-1090 Wien, Austria. E-mail: connyschrauf@chello.at

Received 8 April 2010; revised 16 October 2010; revision accepted 16 October 2010

DOI 10.1002/ajp.20899

Published online 10 November 2010 in Wiley Online Library (wileyonlinelibrary.com).

Visalberghi [2008] replicated Call's experiment with capuchin monkeys (*Cebus apella*), and Schmitt and Fischer [2009] tested olive baboons (*Papio hamadryas anubis*) on the same experimental paradigm. Capuchin monkeys found the food when the provided information derived from shaking the cups (noise or silence produced depending on content). However, the same subjects failed to learn to associate an arbitrary auditory cue with the baited container. Olive baboons were unable to use sound as a cue to find hidden food. Thus, Schmitt and Fischer [2009] restricted the experiment to the use of visual information. First, baboons were shown the content of the empty cup and, based on this information, they had to infer the location of the reward. Six of seven subjects solved this problem. In a second experiment, subjects were given an arbitrary cue (a red cross) as an indicator for the location of the reward. Interestingly, only three of the seven subjects learned to correctly use the arbitrary signal within a given number of trials. This result is even more compelling because when the arbitrary cue experiment was conducted, the subjects already had pre-experience with the underlying method.

Call [2010b] has argued that detecting and exploiting causal relations between stimuli may play a particularly important role in the lives of some species. Two reasons support this conclusion. First, from a theoretical point of view, causal relations are typically grounded on the laws of physics, which means that individuals repeatedly encounter them throughout their lives. Individuals may even be predisposed to interpret events in terms of cause-effect rather than simple co-variation. At the very least, one has to consider that perceiving causal relations may be as basic as perceiving arbitrary relations in terms of co-variation. Second, from an empirical point of view, several studies have established the ubiquity of the causal-arbitrary distinction. Individuals respond differently to even the same stimuli depending on whether they hold a causal or an arbitrary relation [e.g. Call, 2004, 2006a,b; Hanus & Call, 2008; Martin-Ordas & Call, 2009, see Call, 2006b, 2010b for reviews]. This means that the causal-arbitrary distinction does not represent an isolated phenomenon restricted to a specific setup, but it is apparent in a number of situations.

Two aspects of studies on making inferences about food location remain to be investigated. First, there is the issue of the perceptual modality through which information is presented. In previous studies, subjects witnessed an event that produced either a visual or auditory outcome and then selected one of two options available. It is unclear whether subjects could also use kinesthetic information such as weight for such a decision.

Hanus and Call [2008] investigated chimpanzees' notion of weight. More specific, whether chimpanzees can infer the location of food based on

the effect that its weight has on other objects. Chimpanzees were confronted with a balance with two opaque cups mounted on either end. After the experimenter placed a piece of banana in one of the cups behind a screen, subjects saw the balance tipping to one side, and chose the lower cup significantly more often than the higher cup (causal condition). In contrast, subjects chose at chance when the experimenter's action tipped the balance (non-causal balance condition) or when no movement of the balance was involved (non-causal wedge condition). Strikingly, subjects that had been choosing the lower cup in the causal condition reverted to random responses when the weight of the banana was no longer the cause of the movement, demonstrating that chimpanzees had no intrinsic preference for the lower cup. The results once again highlight that paying attention to causal relations was an integral part of the solution to this task.

The ability to use weight information of external objects is not restricted to chimpanzees. Visalberghi and Néel [2003] tested capuchin monkeys' ability to use sound and weight as cues to infer the fullness of a nut. The authors changed the weight of the nuts by leaving the nutshell either fully rewarded (original state), empty, or filled with paper or lead shot. Capuchin monkeys could manipulate both nuts before making a choice. Those subjects touched the nuts and tapped them with their fingers prior to selection, potentially making decisions based not solely on weight but also on sound and smell. Capuchins' ability to use weight information alone has been investigated in the context of nut cracking. Visalberghi et al. [2009] found that capuchins can use the weight of stones or artificial hammers [Schrauf et al. 2008] to decide which hammer may be more effective for cracking open nuts. Furthermore, Klüver [1933] showed that these monkeys can also discriminate objects based on their weight, regardless of their functional value, to solve a task such as nut cracking.

Just like capuchins, great apes can also be trained to discriminate reliably between objects of different weight [McCulloch, 1941; Schrauf & Call, 2009]. McCulloch [1941] trained five chimpanzees to select the heavier of two boxes. Despite the large weight differences (80 vs. 480 g or 80 vs. 640 g), the chimpanzees needed a median of 1,100 trials to master the task. Schrauf and Call [2009] found that orangutans and bonobos could also learn to discriminate objects based on their weight. In that study, subjects had to learn to exchange for a reward one of two identical objects that differed only in weight. Some subjects had to exchange the heavier objects, others the lighter objects. Apes mastered the task but required a median of 331 trials to exceed chance-level performance. Subjects also received an equivalent visual discrimination test in which they had to exchange one of two identical objects that differed

only in achromatic color (black vs. white). Here, subjects needed a median of 64 trials to reach criterion; this is still a relatively large number of trials, especially considering that, prior this study, subjects already had experience with the exchange task. Note, however, that McCulloch's [1941] and Schrauf and Call's [2009] study are both based on solving an arbitrary relation and not a causal one, the correct solution being decided *a priori* by the experimenter.

A second neglected aspect in food location studies is that subjects were never required to seek information themselves about the outcome before making a choice. In all previous studies, subjects were presented with a situation in which they could witness the end result before choosing between the two alternatives, but never produced the outcome themselves. It is therefore important to know whether subjects would play a more active role when the information is not simply provided, i.e. whether they would test and compare different outcomes before making a decision.

Seeking information has been reported in studies devoted to investigating metacognition [e.g. Call, 2010a; Call & Carpenter, 2001; Hampton et al., 2004]. Here, however, the information sought was visual, with on some instances additional auditory information offered [Call, 2010a], and did not necessarily involve making inferences because subjects could directly spy the food. Visalberghi and Néel [2003] also described capuchins seeking information, but it is unclear whether the information was about weight, sound, or perhaps even smell. In the case of nut cracking, Visalberghi et al. [2009] observed capuchins comparing stones for weight and friability but here again, no inference was needed regarding the location of food or any other aspect of the situation.

The current study investigates whether apes are able to seek information for themselves and compare differential outcomes before selecting one of two alternatives when having access only to the containers' weight. We presented apes with two opaque cups and hid slices of two bananas in one of them. Subjects could manipulate the task elements themselves by lifting both cups before making a choice. The rationale for Experiment 1 is to test whether apes, who had already participated in the weight exchange study, can use information about weight to infer the hidden food location. In particular, we were interested to see whether these subjects would independently gain kinesthetic information about cup weight and subsequently select the heavier one.

Experiment 2 was designed to examine whether pre-experience in weight discrimination experiments influenced performance in Experiment 1. We also tested whether apes can learn to use an arbitrary cue to locate the food reward when the weight cue is no longer involved and the causal structure thus

removed. For this purpose, we presented the same task as in Experiment 1 to a group of chimpanzees who were naïve to weight experiments. These subjects also participated in a condition in which cups of the same weight but different visual appearance were presented. The cups were painted with patterns of different color and shape. Subjects had to learn to associate a specific pattern with reward location. Thus, we presented chimpanzees with causal stimuli (i.e. weight indicates the hidden food) and arbitrary stimuli (i.e. a certain visual pattern indicates food), with the contingencies of reinforcement being the same. If apes solved both tasks by establishing arbitrary associations between stimuli (e.g. weight, colors, shapes) and reward, then they should learn the visual cue faster, because a previous study suggests that apes perform better with visual than kinesthetic cues [Schrauf & Call, 2009].

## EXPERIMENT 1

This experiment assessed the general testing procedure and examined whether subjects were capable of using information about weight in order to find hidden food in one of the two cups.

### Methods

#### Subjects

Six orangutans (*Pongo pygmaeus abelii*), five bonobos (*Pan paniscus*), and one gorilla (*Gorilla gorilla gorilla*) housed at the Wolfgang Köhler Primate Research Centre, Leipzig Zoo (Germany), participated in this study. There were eight females and four males ranging from 7 to 34 years of age (Table I). Apart from one orangutan (Kila), all subjects had been tested on weight discrimination prior to this study [Schrauf & Call, 2009] (for a more detailed description, see Table I). All apes lived in social groups with access to indoor and outdoor compounds. Subjects were tested individually in their indoor cages. Feeding took place for all apes species at the same time before testing. All research reported in this manuscript adhered to the American Society of Primatologists (ASP) Principles for the Ethical Treatment of Nonhuman Primates as well as to all German laws regarding animal holding and testing (German "Tierschutzgesetz").

#### Materials

We used two opaque white, conical, paper coffee cups (16.8 cm × 9 cm Ø top side × 6 cm Ø bottom side) with lids (9 cm in diameter) to hide banana slices. The empty and the baited cups weighed 20 and 320 g, respectively. We attached one end of a string to the bottom side of each cup. A small piece of wood, which served as a stopper, was tied to the loose end of the string. A Plexiglas panel with two oval openings

**TABLE I. Species, Name, Age, Sex, Previous Experience in Weight Experiments and Experiment Participation of the Subjects**

Species	Subject	Age (years)	Sex	Previous experience in weight experiments	Experiment participation
Orangutan	Padana	10	F	Pass	1
Orangutan	Dokana	18	F	Fail	1
Orangutan	Pini	19	F	Pass	1
Orangutan	Dunja	34	F	Fail	1
Orangutan	Bimbo	27	M	Fail	1
Orangutan	Kila	7	F	Not tested	1
Bonobo	Joey	25	M	Fail	1
Bonobo	Kuno	11	M	Pass	1
Bonobo	Limbuko	12	M	Pass	1
Bonobo	Ulindi	14	F	Pass	1
Bonobo	Yasa	10	F	Fail	1
Gorilla	Viringika	12	F	Fail	1
Chimpanzee	Alex	7	M	Not tested	2
Chimpanzee	Annette	9	F	Not tested	2
Chimpanzee	Alexandra	9	F	Not tested	2
Chimpanzee	Jahaga	15	F	Not tested	2
Chimpanzee	Gertruida	15	F	Not tested	2
Chimpanzee	Fifi	15	F	Not tested	2
Chimpanzee	Pia	8	F	Not tested	2
Chimpanzee	Fraukje	32	F	Not tested	2
Chimpanzee	Lome	6	M	Not tested	2

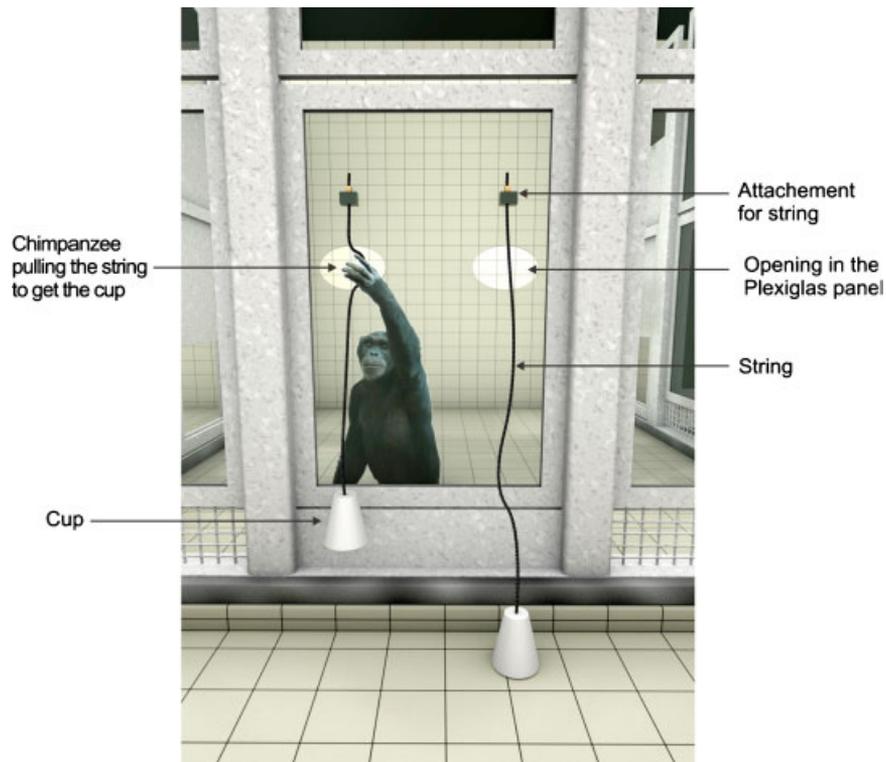


Fig. 1. Experimental setup of the testing apparatus: Subject pulls the string to lift the cup towards the opening. The cups were visually identical but differed in weight. Drawing by Sylvio Tuepke.

(11.5 cm × 8.5 cm; 40 cm apart) on its upper part served for the presentation of the two cups (Fig. 1). A small plastic block (5 × 2.3 × 4 cm) was attached to

the Plexiglas above each opening, and the string passed through a slit in the block. The subjects could grasp the string through the opening in the panel.

Pulling the string lifted the upside-down cup upwards along the panel and, as soon as it reached the opening, the subjects could touch the cup with their hands. The opening size allowed the apes to grasp the string but prevented them from grabbing the cup. A gray PVC box with an opening at the back ( $82 \times 38 \times 30$  cm) served to conceal the shuffling of the cups from the subject.

### Procedure

The experimenter (E) sat outside the cage and waited until the subject sat in front of the panel to start the trial. First, the E showed the two empty cups to the ape. Then, the E presented a bucket filled with banana slices. E completely filled one of the cups in full view of the subject, whereas the other cup remained empty. The E showed the subject the empty bucket, covered the cups with lids and then shuffled them behind the occluder. Finally, the occluder was removed and the strings were passed through the block on the Plexiglas panel. To indicate their choice, apes had to touch one of the cups with their fingers. To do so, they had to grasp the string and pull it until the cup reached the opening. Only if the subject touched the cup we counted it as choice. Pulling the string and the cup halfway and then releasing it was not counted as a choice. Since holding the string and touching the cup required using both hands, the apes' choices were unambiguous. After a subject made a choice, we unhooked the strings and showed the ape the content of both cups. If the subject had selected the baited cup, it received its content. Otherwise the E put the content of the cup back into the bucket and left the testing room. Note that we did not familiarize our subjects with the procedure prior to testing as we were interested in apes' spontaneous behavior. In particular, whether they would independently seek kinesthetic information about the cups weight before making a choice. The ordering of side of food presentation was determined randomly, with the only constraint that it was not hidden in the same location in more than two successive trials. We conducted 16 trials on consecutive days with eight trials per side for each subject. Due to the large amount of food (two bananas per trial), we conducted only 1 trial per day.

### Data scoring and analysis

All trials were videotaped. We used two-tailed non-parametric statistics to analyze the data. Our main dependent variable was the number of correct trials (e.g. the number of trials a subject selected the heavy cup). We compared the performance of bonobos and orangutans with the Mann-Whitney test. We used the Wilcoxon test to compare subjects' performance as a group to chance for all and for only the second half of trials separately. Furthermore, the Wilcoxon test was used to analyze whether subjects

improved their performance between the first eight trials and the last eight trials. We used the binomial test (expected  $P = 0.5$ ) to analyze the performance on the first trial for the entire group and to test the individual performance over all 16 trials. Specifically, whether a subject had selected the correct cup at above chance level, that is in at least 12 of 16 trials. We used Fisher's test to assess the effect that a previous experiment on weight discrimination had on the current study. We also scored whether subjects compared between the two cups by successfully pulling both strings before making a choice. We used Spearman correlation to investigate the relation between success and pulling both strings in each trial. Finally, we used Wilcoxon tests to compare the frequency of pulling the correct and then incorrect cup with the frequency of pulling the incorrect and then the correct cup.

## Results

### Success

Bonobos performed significantly better than orangutans (Mann-Whitney  $U$  test  $U = 4$ ;  $N_1 = 5$ ;  $N_2 = 6$ ;  $P = 0.05$ ; mean (SEM): bonobos = 11.8 (SEM = 1); orangutans = 8.3 (SEM = 1.6)). All bonobos except one performed above chance levels, whereas only two of the six orangutans did so (Binomial test:  $P < 0.05$ , Table II). Subjects as a group did not perform above chance levels (Wilcoxon exact test  $Z = 1.596$ ;  $N = 12$ ;  $P = 0.121$ ). Errors occurred mainly in the first half of the experiment (Trial 1–8) compared with the second half (Trial 9–16) (Wilcoxon exact test  $Z = 2.223$ ;  $N = 12$ ;  $P = 0.03$ ). Reanalyzing the data for only the second half of the trials revealed that subjects as a group performed above chance levels (Wilcoxon exact test  $Z = 2.103$ ;  $N = 12$ ;  $P = 0.031$ ). An analysis of the first trial performance showed no initial preference for the heavy cup (Binomial test  $P = 0.388$ ). The type of positive stimulus (light/heavy) that had been associated with reward in the Schrauf and Call [2009] weight discrimination study did not designate success in the present study (Fisher exact test  $P = 0.43$ ). Additionally, success in the former experiment did not determine success in the current one (Fisher exact test  $P = 0.16$ ).

### Pulling patterns

Figure 2 presents the percent of successful trials as a function of the number of trials in which subjects successively pulled at both strings. The more subjects pulled on both strings, the more likely they were to succeed in finding the food (Spearman correlation  $r(12) = 0.83$ ,  $P = 0.001$ ). In the first trial, however, only Pini pulled both strings, and ended up selecting the empty cup.

Figure 3 presents the mean number of trials in which subjects performed a particular pulling pattern. Most responses involved a single pull that

TABLE II. Participants of Experiment 1

Species	Subject	Performance	Weight exchange	
			Correct stimulus	Performance
Orangutan	Padana	<b>12*</b>	Light	Pass
Orangutan	Pini	7	Light	Pass
Orangutan	Dokana	<b>12*</b>	Heavy	Fail
Orangutan	Dunja	11	Heavy	Fail
Orangutan	Bimbo	5	Light	Fail
Orangutan	Kila	3		Not tested
Bonobo	Joey	8	Heavy	Fail
Bonobo	Limbuko	<b>13*</b>	Light	Pass
Bonobo	Kuno	<b>12*</b>	Light	Pass
Bonobo	Yasa	<b>13*</b>	Heavy	Fail
Bonobo	Ulindi	<b>13*</b>	Heavy	Pass
Gorilla	Viringika	8	Heavy	Fail

Numbers in bold indicate successful performance. Weight exchange correct: type of correct stimulus in the exchange experiment. Weight exchange performance: Successful: subjects reached required criterion. Failure: subjects did not meet criterion.

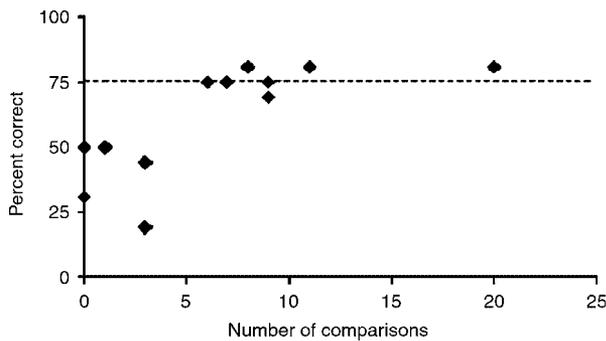


Fig. 2. Number of comparisons over the course of the experiment in relation to success for all subjects in Experiment 1. Primary axis: Number of trials in which subjects successively pulled on both strings. All pulls are considered, independent of whether subjects ended up with the correct or incorrect cup. Secondary axis: Percent of correct trials. Dashed line indicates above chance performance ( $P < 0.05$ ).

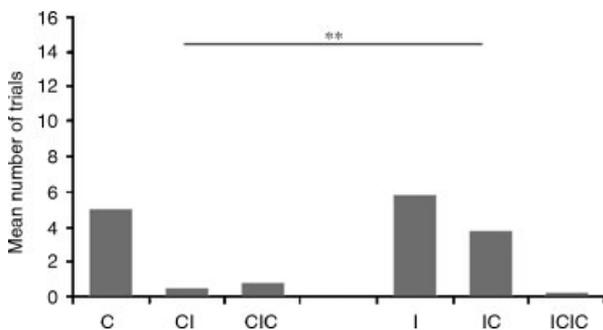


Fig. 3. Mean number of times subjects showed the following behavior: C = correct I = incorrect C or I = single pull; CI or IC = 1 comparison; CIC = 2 comparisons; ICIC = 3 comparisons; First letter indicates the cup subjects started to pull; Last letter indicates the cup subjects ended up with.

resulted in a correct or incorrect choice. Moreover, there was no difference in the number of times apes chose the correct or incorrect cup after their first pull (Wilcoxon exact test  $Z = 0.359$ ;  $P = 0.773$ ). However,

when subjects switched after their first pull, they did so more often from the incorrect cup to the correct cup (IC) than vice versa (CI) (Wilcoxon exact test  $Z = 2.673$ ;  $P = 0.004$ ). The remaining cases depicted in Figure 3 were too few to enable statistical analyses, but in every case the apes ended up selecting the correct cup. Kila, the orangutan who had not participated in the previous weight experiment, rarely switched from one cup to the other. In particular, she switched twice from the correct to the incorrect cup (Trial 4, 13) and once from the incorrect to the correct cup (Trial 15). Viringika, the gorilla female, never switched between cups.

## DISCUSSION

Six of the 12 subjects tested succeeded in using the weight of the cup to find the hidden food. Apparently, the essential step to solve the task consisted in gaining kinesthetic information about the weight of each of the cups and subsequently selecting the heavier one. Successful apes spontaneously adopted the method of successive comparisons of the cups' weight. Compared with the results of Schrauf and Call [2009], subjects showed rapid improvement in learning to discriminate weight even though the weight differences were the same in both studies. This result is even more striking because we used a novel method for the subjects in the present study, whereas in the earlier exchange study subjects were already familiar with the exchange procedure prior to testing. In the current study, subjects might have learned to use weight more easily because of the causal structure of the task, i.e. food causes weight.

Conceivably, the apes' good performance was influenced by their previous experience with weight experiments, although we detected no significant effect of their previous performance on the current results. Having been tested on weight discrimination before the current experiment may have made

weight a more relevant dimension to them. Furthermore, one could argue that the fast learning was related to the large amount of food they received as reward. To test these alternative explanations, we ran a second experiment with subjects that had no previous experience on weight discrimination. Additionally, we introduced a second condition that, like the Schrauf and Call [2009] study, was based on an arbitrary (yet equally predictive) relation between the stimuli and the reward. Here, the causal structure was removed as an arbitrary visual cue indicated the location of the reward. If the abundant food promoted the fast learning, and if the task was solved by establishing arbitrary associations between stimuli and the presence of the reward, then the apes should learn the visual cue faster because previous studies suggest that they perform better with visual than weight cues.

## EXPERIMENT 2

In this experiment, we examined the performance of naïve chimpanzees in finding hidden food by using its weight (causal condition), just like in Experiment 1. We also investigated whether they could solve the task by using visual information (i.e. color and shape of stickers) as discriminative cues (arbitrary condition) when such cues were not produced by the presence of food. The reason for adding visual information to weight was because previous studies had suggested greater proficiency in visual compared with weight discriminations.

### Methods

#### Subjects

Nine chimpanzees (*Pan troglodytes*) housed at the Wolfgang Köhler Primate Research Centre, Leipzig Zoo (Germany) participated in this study. The seven females and two males ranged from 6 to 32

years of age (Table I). All subjects were naïve to weight experiments and were tested individually in their indoor cages. The subjects were never food deprived during the experiment and water was available ad libitum.

#### Materials

We used the same basic apparatus as in Experiment 1 including the two cups, lids and attached strings. For the causal condition, we used two white cups, whereas in the arbitrary condition we used two olive green cups, which differed in the color (green or blue) and shapes (circles vs. triangles) of the stickers that were attached to them (see Table III for information on the precise patterns used). In the causal condition, the empty and the baited cups weighed 20 and 320 g, respectively. In the arbitrary condition, both cups weighed the same after the baiting, because the empty cup was filled with a mixture of lead shot and silicone. Therefore, in the arbitrary condition, the weight cue was no longer involved. The cups in the arbitrary condition could only be discriminated by their visual appearance. To hide the lead shot inside the cup from the subject, we glued paper on top of it. Both cups therefore looked identical when they were empty and shown to the subjects. The solid mass formed by the silicone and lead shot distributed the weight evenly in the cup and prevented any rattling noise during manipulations.

#### Procedure

There were two conditions: causal and arbitrary. The order of condition was counterbalanced across subjects. The general procedure was identical to that of Experiment 1. The only changes refer to the arbitrary condition. To make the two conditions as similar as possible, we placed the cups in the arbitrary condition in a gray PVC box

TABLE III. Participants, Pairing of Presented Cups and Correct Stimulus of Experiment 2

Subject	Pair of presented cups	Correct stimulus	Condition		Order
			Arbitrary condition	Causal condition	
Pia	Green ○ vs. yellow Δ	Green ○	8	9	c-a
Fraukje	Blue Δ vs. green ○	Blue Δ	7	10	c-a
Lome	Blue Δ vs. green ○	Blue Δ	8	8	c-a
Jahaga	Blue Δ vs. green ○	Blue Δ	5	<b>16</b>	a-c
Alexandra	Yellow Δ vs. blue ○	Yellow Δ	7	<b>14</b>	a-c
Fifi	Yellow Δ vs. blue ○	Yellow Δ	6	<b>12</b>	c-a
Anette	Green ○ vs. blue Δ	Green ○	6	<b>15</b>	a-c
Gertruida	Green ○ vs. blue Δ	Green ○	8	6	a-c
Alex	Blue Δ vs. green ○	Blue Δ	<b>13</b>	<b>15</b>	c-a
		Mean	7.6	11.7	
		% correct	48	73	

Numerals represent in how many trials subjects were correct in the respective condition. Numbers in bold indicate successful performance. Note that a minimum of 12/16 is necessary to perform above chance level. Order: order in which the two tests were performed (c-a: subjects were first tested in the causal and thereafter in the arbitrary condition, a-c: vice versa)

(82 × 8.4 × 11 cm) before starting a trial. Without lifting the cups, subjects were unable to see the stimuli that were fixed on the lower half of the cup. Thus, subjects had to lift the cup a bit in order to perceive the stimuli. Each chimpanzee received 16 trials per condition on consecutive days. All trials of one type had to be completed before subjects received any trials of the other type.

#### Data scoring and analysis

The same scoring method and analyses as in Experiment 1 was used for the causal and arbitrary condition in Experiment 2. Additionally, we compared the performance in the causal condition with the arbitrary condition using the Wilcoxon test. To test for species differences, we compared the performance of bonobos, orangutans (Experiment 1) and chimpanzees (Experiment 2 causal condition) using the Kruskal–Wallis test.

## Results

### Success

Subjects performed significantly better in the causal than arbitrary condition (Wilcoxon exact test  $Z = 3.728$ ;  $N = 18$ ;  $P < 0.001$ ). Moreover, subjects performed above chance level in the causal (Wilcoxon exact test  $Z = 2.176$ ;  $N = 9$ ;  $P = 0.031$ ) but not in the arbitrary condition (Wilcoxon exact test  $Z = 0.949$ ;  $N = 9$ ;  $P = 0.375$ ). Focusing on the second half of the trials produced the same results (Wilcoxon exact tests: causal:  $Z = 2.345$ ;  $N = 9$ ;  $P = 0.023$ ; arbitrary:  $Z = 0.431$ ;  $N = 9$ ;  $P = 0.781$ ).

Five of the nine chimpanzees performed above chance levels in the causal condition (Binomial test:  $P < 0.05$ , Table III), whereas only one did so in the arbitrary condition (Binomial test:  $P < 0.05$ , Table III). The order of presentation of conditions had no influence on performance in the causal (Fisher exact test:  $P = 0.318$ ) and arbitrary condition (Fisher exact test:  $P = 0.556$ ).

Comparing the number of errors in the first half (Trial 1–8) and second half (Trial 9–16) of the experiment showed no difference for the causal (Wilcoxon exact test:  $Z = 1.474$ ;  $N = 9$ ;  $P = 0.17$ ) or arbitrary condition (Wilcoxon exact test  $Z = 0.954$ ;  $N = 9$ ;  $P = 0.438$ ). In the former, however, only one of the successful subjects (Fifi) made errors in the second half of the study (Trial 15, 16)—and she did so after being correct on 10 successive trials. An analysis of the first trial performance showed no initial preference for either the light or heavy cup (Binomial test:  $P = 1$ ).

Combining the data from Experiments 1 and 2 (causal condition) showed no significant differences between species (Kruskal–Wallis test:  $\chi^2 = 3.97$ ;  $N_1 = 5$ ;  $N_2 = 6$ ;  $N_3 = 9$ ;  $P = 0.137$ ; mean (SEM): bonobos = 11.8 (SEM = 1), orangutans = 8.3 (SEM = 1.6), and chimpanzees = 11.7 (SEM = 1.2).

### Pulling patterns

Figure 4A and B shows the percent of successful trials as a function of the number of trials in which both strings were pulled successively. The more both strings were pulled, the greater the success in finding the food both in the causal (Spearman correlation  $r(9) = 0.755$ ,  $P = 0.019$ ) and arbitrary conditions (Spearman correlation  $r(9) = 0.777$ ,  $P = 0.014$ ). In the causal condition, four subjects compared the cups on their first trial and three of them ending up with the correct cup. In the arbitrary condition, two subjects compared the cups on their first trial, both ending up with the correct cup.

Figure 5A presents the mean number of trials in which subjects performed a particular pulling pattern in the causal condition. In those cases involving a single pull (C & I), there was no difference in the number of times apes chose the correct or incorrect cup (Wilcoxon exact test  $Z = 1.577$ ;  $P = 0.125$ ). However, when subjects switched after their first pull, they did so more often from the incorrect to the correct cup (IC) than vice versa (CI) (Wilcoxon exact test  $Z = 2.388$ ;  $P = 0.016$ ). With three exceptions, on all remaining occasions in which subjects compared the cups more often (two to five times), they always ended up selecting the correct cup.

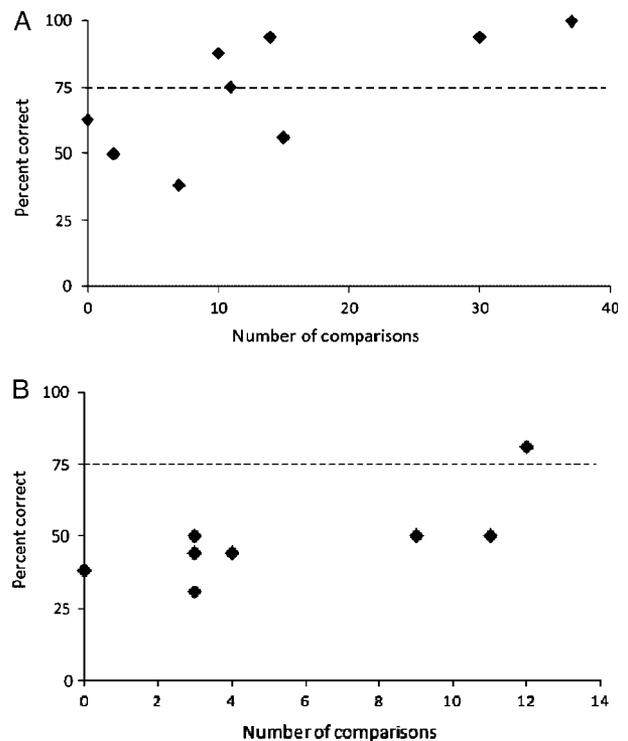


Fig. 4. (A) Causal condition; (B) Arbitrary condition. Number of comparisons over the course of the experiment in relation to success for all subjects in Experiment 2. Primary axis: Number of trials in which subjects successively pulled on both strings. All pulls are considered, independent of whether subjects ended up with the correct or incorrect cup. Secondary axis: Percent of correct trials. Dashed line indicates above chance performance ( $P < 0.05$ ).

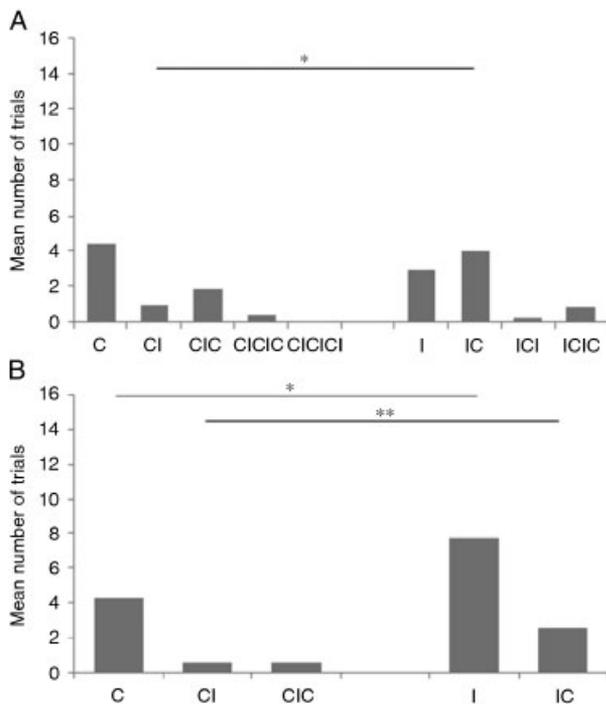


Fig. 5. (A) Causal condition; (B) Arbitrary condition. Mean number of times subjects showed the following behavior: C = correct I = incorrect C or I = single pull; CI or IC = 1 comparison; CIC = 2 comparisons; ICIC = 3 comparisons; CICIC = 4 comparisons; CICICI = 5 comparisons. First letter indicates the cup subjects started to pull; Last letter indicates the cup subjects ended up with.

Figure 5B shows the mean number of trials in which a certain pulling pattern was performed in the arbitrary condition. In single pulls (C & I), apes chose the incorrect cup more often than the correct cup (Wilcoxon exact test  $Z = 2.673$ ;  $P = 0.004$ ). When they switched after their first pull, they did so more often from the incorrect to the correct cup (IC) than vice versa (CI) (Wilcoxon exact test  $Z = 2.226$ ;  $P = 0.031$ ). In the remaining cases in which subjects switched between cups, they ultimately chose the correct cup.

## Discussion

Five of the nine chimpanzees tested were able to find the food in the causal condition (i.e. food causes weight). In contrast, only one subject solved the task in the arbitrary condition, where visual cues indicated the correct cup. Interestingly, in the causal condition naïve subjects showed a higher frequency of multiple comparisons per trial before making a choice than experienced subjects in Experiment 1. Moreover, whereas experienced subjects showed a maximum of three comparisons per trial, naïve subjects compared the cups four and even five times per trial. The effect was opposite for the arbitrary condition. In particular, multiple comparisons per trial occurred less often in the arbitrary versus

causal condition as well as compared with Experiment 1. It is important to note that lifting the cups in the arbitrary condition was necessary to have visual access to the stimuli but not to receive kinesthetic feedback. Our findings show that all apes react to weight cues, both those with experience in weight discrimination and naïve ones. And most importantly, that even the naïve subjects spontaneously sought kinesthetic information for themselves before making a choice. Naïve subjects were less sensitive to weight differences and thus compared more than experienced subjects. Subjects' good performance in the causal condition corroborated our findings from Experiment 1 and clearly showed that the success in the causal condition cannot be explained by previous experience on weight experiments. Furthermore, when the weight cue and thus the causal structure was removed, subjects were unable to select the correct cup above chance levels even after 16 trials. Despite the visual cues, only one subject was able to solve the task in the arbitrary condition. This makes it unlikely that abundant food accounts for the fast speed of learning in the causal condition of both experiments.

## General Discussion

Great apes were able to find hidden food in one of two opaque cups using information about the food's weight. Success in this task required seeking kinesthetic information about the weight of each of the cups and subsequently selecting the heavier one. Apes spontaneously adopted the method of successively pulling the strings and thus comparing weights before choosing. Unlike what we observed in Schrauf and Call [2009], subjects showed rapid improvement in learning to discriminate weight although the weight differences were the same in both studies. Although subjects needed a median of 331 trials until mastery in the weight discrimination task [Schrauf & Call, 2009], here they reached above chance level within 16 trials. This result is even more remarkable because, when the exchange task in the earlier study was conducted, the subjects already had pre-experience with the underlying method, the only difference being the tested feature (weight). In contrast, the current study involved a method that was novel for the subjects.

We presented task elements in which the stimuli and the reward had a cause-effect relation (i.e. food causes the cue) and task elements in which no causal relation existed between elements (i.e. food co-varies with the cue). Only one subject was able to solve the problem when the task elements involved arbitrary relations. Subjects that began with the causal condition and had successfully solved the task reverted back to chance level in the arbitrary condition. This result is congruent with Hanus and Call's [2008] findings on the balance beam problem.

In their study, chimpanzees that had chosen the lower cup in the causal condition (when food caused the movement) reverted back to chance when the causal structure was removed and arbitrary stimuli served as cue for the location of food.

The difference in performance (causal versus arbitrary condition) and the speed of learning between the causal condition and the earlier weight discrimination experiment [Schrauf & Call, 2009] can be explained by the structure of the presented tasks. The aspect of arbitrary and nonarbitrary relations between the cue and the food plays a major role in performance.

Although both types of cues possess the same predictive value, namely 100% for the correct reward location, only one of them has causal relevance to the task at hand. In the present study, both weight and color/shape are indicators for the presence of food. Nevertheless, only the weight cue was caused by the food and thus provided a source of information that is grounded on the laws of physics. The appearance of a certain color as predictor for the presence of food is merely based on co-variation. Similarly, in the earlier weight discrimination experiment no causal relation existed between the task elements. The experimenter decided which stimulus (heavy/light) was correct for each subject. Thus, they had to learn through a history of reinforcement to associate a particular stimulus with the reward.

The causal-arbitrary distinction is a robust phenomenon found in various studies that used different modalities, e.g. visual or auditory [Call, 2006a; Hanus & Call, 2008; Sabbatini & Visalberghi, 2008]. In all described cases, individuals never attended to the cue (orientation, shape, and noise) *per se* as indicator for the food location, but their responses critically depended on whether the hidden food caused or simply co-varied with the observable cue. For example, when searching for food, apes prefer to select a slanted over a flat board when the slanted orientation is caused by the presence of the food underneath. When a wedge serves as an indicator for the food, however, apes do not prefer the slanted over the flat board [Call, 2007]. Similarly, in the auditory modality, apes show a clear difference in performance when using causal vs. arbitrary information to find hidden food. In particular, subjects use noise as a source of information when it is caused by the food inside a cup that is shaken. In contrast, subjects fail to use the same noise when it is produced by a tape recorder positioned on the corresponding cup.

With regard to our current results, one could argue that animals had learned from previous experience that heavier cups contained food and light ones did not. Such preference for heavier cups would have given them an advantage in the weight compared to the visual condition. However, such putative advantage is not supported by other data.

Not only weight discrimination in the absence of a causal structure is hard [see McCulloch, 1941; Schrauf & Call, 2009] but our subjects did not show an intrinsic preference for heavier containers over lighter ones [Schrauf & Call, 2009], something that would have been expected if there was a history of learning biasing the results of this study. Additionally, in terms of learning to associate cues with the presence of food, our subjects had received a number of visual discrimination tests but only one on weight discrimination. Based on the learning-set phenomenon [Harlow, 1949; see also Tomasello & Call, 1997], one would expect that subjects should perform better with visual than weight cues.

Another possible explanation for our results is that subjects have a stronger predisposition (independently from learning) to attend to weight rather than color/shape cues. However, this hypothesis is considerably weakened by some of the data available in the literature. For instance, apes can use the shape of an object very effectively when the shape is caused by the presence of food, but they do not use it when it simply co-varies with the presence of food [see Braeuer et al., 2006; Call, 2007]. We propose that individuals are not prepared to attend to weight or color/shape *per se*, but that their responses critically depend on the relation between the cue and the reward.

We can rule out the possibility that performance differences between the causal and the arbitrary condition are because of weight being a more salient cue than color or shape. On the contrary, previous studies have shown that apes face difficulties when required to discriminate weight. As noted earlier, McCulloch's [1941] chimpanzees needed about 1,100 trials to master the task, and our own study on weight discrimination using the exchange paradigm required 331 trials. In fact, Schrauf and Call [2009] showed that subjects learned color discrimination faster than weight discrimination. Moreover, if one of the conditions in this study had a more salient cue, then it would have been the arbitrary condition because individuals could rely on a certain color and/or shape. We can exclude that the abundant reward in the present study promoted learning: subjects in the arbitrary condition received the same amount of food, and still only one subject was able to solve the task. Finally, we can rule out the possibility that fast learning reflects previous experience in weight experiments because we ran the same experiment with a group of naïve chimpanzees. Even though these subjects had not been tested on weight experiments in the past, they also performed above chance level within a total of 16 trials. Note also that half of the participants in the weight conditional discrimination task had learned to respond to light weight as the type of positive stimulus.

One could argue that some studies have failed to find evidence that primates are sensitive to cues that are causally relevant as opposed to arbitrary. For

example, Vonk and Subiaul [2009] tested chimpanzees' ability to predict which of two human experimenters could deliver a tray containing food based on the visibility or the freedom of movement of the experimenter's limbs. For instance, chimpanzees failed to select the experimenter whose legs were not blocked by an obstacle. The authors concluded that chimpanzees focused on cues that were functionally irrelevant but ignored cues that were functionally relevant. However, the problem that Vonk and Subiaul [2009] tackled is different from ours. First, we explored how apes perceive object-object relations, whereas Vonk and Subiaul [2009] investigated how chimpanzees perceived the potential for human action. Second, in our study, both weight and color invariably predicted the presence of food regardless of whether the cue held a causal as opposed to an arbitrary relation with the presence of food. Hence the critical feature that we were contrasting was not "relevant" vs. "irrelevant" cues but whether apes responded differently to relevant cues that differed in the relation they held with the reward.

Neither subjects with weight learning experience nor naïve subjects showed an initial preference for the heavy or light cup in this task. After a few trials, successful apes introduced the method of successively pulling the strings before making a choice. Note that we did not present the weights consecutively, but the apes themselves found a way to compare the cups in sequence. The essential step to success seems to involve seeking information about the weight of both cups by successively comparing them before choosing.

Naïve subjects (Experiment 2), however, showed a higher frequency of multiple comparisons per trial, before making a choice, than experienced subjects (Experiment 1). Moreover, whereas experienced subjects showed a maximum of three comparisons per trial, naïve subjects compared the cups four and even five times per trial. A possible explanation for this difference is that experienced subjects had an advantage by being already more sensitive to the weight differences. In fact, when Klüver decreased the weight differences in his experiments from 450:150 to 325:250 g, the number of comparisons subjects did per trial increased. He explained these findings by arguing that "the great numbers of comparisons were undoubtedly indicative of the great uncertainty of the monkeys (p 62)." This explanation is further supported by the fact that, in the arbitrary condition, multiple comparisons per trial occurred less often compared to the causal condition in Experiment 2 vs. Experiment 1. Note that, in the arbitrary condition, lifting the cups was not important to perceive the kinesthetic cue. Instead, it was necessary to have access to the visual pattern on the cups that predicted the food location. The fact that subjects compared less often in the

arbitrary compared to the causal condition might mean that they were less sensitive to kinesthetic than visual cues. This assumption is further strengthened by previous findings that apes learn visual cues faster than weight cues [compare McCulloch, 1941 and Nissen & Jenkins, 1943; see also Schrauf & Call, 2009].

Another possibility could be that apes do not learn weight in an absolute sense (heavy/light) but in a relative sense (heavier/lighter than). In the earlier weight exchange task [Schrauf & Call, 2009], successful subjects separated the objects spatially according to whether they were heavy or light. Because an object's weight is only perceived by lifting it, this behavior could have helped the subjects to keep track of the correct and incorrect objects. This strategy could likewise result from the subjects' difficulty to learn the absolute weight of an object. Furthermore, in the present study, the high number of comparisons of the cups weight before making a choice could also indicate a necessity to set the perceived weights in relation before selecting the heavier one.

The ability to use the weight of external objects as a source of information has been described so far only for capuchin monkeys [Schrauf et al., 2008; Visalberghi et al., 2009] and chimpanzees [Bril et al., 2009; Hanus & Call, 2008]. We provide the first evidence of this capacity in bonobos and orangutans. Although we did find species differences, with bonobos performing better than orangutans, we urge caution when interpreting this finding. Due to the small sample size, performance discrepancies can likewise reflect individual instead of species differences. To the best of our knowledge, there is no good ecological reason that would explain why bonobos should be better at using weight cues than orangutans. For those species exploiting hard-shelled nuts, it would be advantageous to first assess whether the kernel is present by the nuts' weight. However, neither bonobos nor orangutans have been documented exploiting nuts in their natural habitats. The degree of arboreality could be another potential explanation. The problem is that given the higher degree of arboreality of orangutans compared to bonobos, and orangutans' specialized means of locomotion in the canopy [see Povinelli and Cant, 1995], one would expect the differences to appear in the opposite direction than observed, that is, orangutans outperforming bonobos. However, note that our study investigated the effect of an external object on another object not the subject's own weight on another object, which is perhaps where orangutans may outperform other terrestrial apes.

Chimpanzees performed equally well as bonobos and orangutans if all species are analyzed together. One caveat of this comparison is that the latter two, unlike chimpanzees, had previously participated in a study that involved discriminating two weights

[Schrauf & Call, 2009]. Although we detected no effect of the treatment that bonobos and orangutans had received on our current data, we cannot rule out the possibility that this experience affected them in some way, thus compromising the comparison with the experimentally naïve (with regard to weight) chimpanzees. Future studies should test naïve bonobos and orangutans on our task. We did not do this ourselves because we had no access to naïve individuals.

In summary, our findings show that apes quickly learned to use weight information in a situation in which weight held a causal relation with the task solution. In contrast, subjects failed to solve the same problem when it was based on learning to associate arbitrary visual stimuli with the presence of the reward. Our results confirm previous findings [Call, 2004, 2006a,b, 2007; Hanus & Call, 2008] that apes and monkeys [Sabbatini & Visalberghi, 2008; Schmitt & Fischer, 2009] perform better in problems grounded on causal compared to arbitrary relations. Future studies are needed to determine how widespread this ability is across other species.

## ACKNOWLEDGMENTS

This research was supported by a grant (DOC fFORTE) from the Austrian Academy of Science to Cornelia Schrauf. We thank Sylvio Tuepke for his drawing of the experimental setup. The reported experiments comply with all laws of the country in which they were performed.

## REFERENCES

- Blaisdell AP, Sawa K, Leising KJ, Waldmann MR. 2006. Causal reasoning in rats. *Science* 311:1020–1022.
- Braeuer J, Kaminski J, Riedl J, Call J, Tomasello M. 2006. Making inferences about the location of hidden food: social dog-causal ape. *Journal of Comparative Psychology* 120:38–47.
- Bril B, Dietrich G, Foucart J, Fuwa K, Hirata S. 2009. Tool use as a way to assess cognition: how do captive chimpanzees handle the weight of the hammer when cracking a nut? *Animal Cognition* 12:217–235.
- Call J. 2004. Inferences about the location of food in the Great Apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, and *Pongo pygmaeus*). *Journal of Comparative Psychology* 115:159–171.
- Call J. 2006a. Descartes' two errors: Reason and reflection in the great apes. In: Hurley S, Nudds M, editors. *Rational animals*. Oxford: Oxford University Press. p 219–234.
- Call J. 2006b. Inferences by exclusion in the great apes: the effect of age and species. *Animal Cognition* 9:393–403.
- Call J. 2007. Apes know that hidden objects can affect the orientation of other objects. *Cognition* 105:1–25.
- Call J. 2010a. Do apes know that they could be wrong? *Animal Cognition* 13:689–700.
- Call J. 2010b. Understanding apes to understand humans: the case of object-object relations. In: Glatzeder BM, Goel V, von Muller A, editors. *Towards a theory of thinking*. Heidelberg: Springer. p 215–230.
- Call J, Carpenter M. 2001. Do apes and children know what they have seen? *Animal Cognition* 3:207–220.
- Hampton RR, Zivin A, Murray EA. 2004. Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Animal Cognition* 7:239–254.
- Hanus D, Call J. 2008. Chimpanzees infer the location of a reward based on the effect of its weight. *Current Biology* 18:R370–R372.
- Harlow HF. 1949. The formation of learning sets. *Psychological Review* 56:51–65.
- Klüver H. 1933. *Behavior mechanisms in monkeys*. Chicago: Chicago University Press.
- Martin-Ordas G, Call J. 2009. Assessing generalization within and between trap tasks in the great apes. *International Journal of Comparative Psychology* 22:43–60.
- McCulloch TL. 1941. Discrimination of lifted weights by chimpanzees. *Journal of Comparative Psychology* 32:507–519.
- Nissen HW, Jenkins WO. 1943. Reduction and rivalry of cues in the discrimination behavior of chimpanzees. *Journal of Comparative Psychology* 35:85–95.
- Povinelli DJ. 2000. *Folk physics for apes: the chimpanzee's theory of how the world works*. Oxford: Oxford University Press.
- Povinelli DJ, Cant JGH. 1995. Arboreal clambering and the evolution of self-conception. *Quarterly Review of Biology* 70:393–421.
- Sabbatini G, Visalberghi E. 2008. Inferences about the location of food in capuchin monkeys (*Cebus apella*) in two sensory modalities. *Journal of Comparative Psychology* 122:156–166.
- Schmitt V, Fischer J. 2009. Inferential reasoning and modality dependent discrimination learning in olive baboons (*Papio hamadryas anubis*). *Journal of Comparative Psychology* 123:316–325.
- Schrauf C, Call J. 2009. Great apes' performance in discriminating weight and achromatic color. *Animal Cognition* 12:567–574.
- Schrauf C, Huber L, Visalberghi E. 2008. Capuchin monkeys learn to use weight to select the most effective tool to crack open nuts. *Animal Cognition* 11:413–422.
- Seed AM, Call J. 2009. Causal knowledge for events and objects in animals. In: Watanabe S, Blaisdell AP, Young A, editors. *Rational animals, irrational humans*. Tokyo: Keio University Press. p 173–187.
- Seed AM, Tebbich S, Emery NJ, Clayton NS. 2006. Investigating physical cognition in Rooks, *Corvus frugilegus*. *Current Biology* 16:697–701.
- Thorndike EL. 1898. *Animal intelligence. An experimental study of the associative processes in animals*. *Psychological Review Monograph Supplements* 2:551–553.
- Tomasello M, Call J. 1997. *Primate cognition*. Oxford: Oxford University Press.
- Visalberghi E, Néel C. 2003. Tufted capuchins (*Cebus apella*) use weight and sound to choose between full and empty nuts. *Ecological Psychology* 15:215–228.
- Visalberghi E, Addressi E, Truppa V, Spagnoletti N, Ottoni E, Izar P, Frigaszy D. 2009. Selection of effective stone tools by wild capuchin monkeys. *Current Biology* 19:213–217.
- Vonk J, Subiaul F. 2009. Do chimpanzees know what others can and cannot do? *Animal Cognition* 12:267–286.
- Woodward J. 2007. Interventionist theories of causation in psychological perspective. In: Gopnik A, Schultz L, editors. *Causal learning: psychology, philosophy and computation*. Oxford: Oxford University Press. p 16–36.