The Magic Cup: Great Apes and Domestic Dogs (*Canis familiaris*) Individuate Objects According to Their Properties

Juliane Bräuer and Josep Call
Max Planck Institute for Evolutionary Anthropology

Despite current interest in dog (*Canis familiaris*) cognition, very little is known about how dogs represent objects and how they compare with other species, such as the great apes. Therefore, we investigated how dogs and great apes (chimpanzees [*Pan troglodytes*], bonobos [*Pan paniscus*], orangutans [*Pongo pygmaeus*], gorillas [*Gorilla gorilla*]) individuate objects in a classical violation of expectation paradigm. We used a container (magic cup) with a double bottom that allowed us to change the type of food that subjects had seen being placed in the container. Using a 2 × 2 design, we varied whether subjects received a generally preferred food and whether the food was substituted (surprise trials) or not (baseline trials). Apes showed increased begging and looking behaviors and dogs showed increased smelling behavior. Both species stayed near the experimenter more frequently in the surprise trials compared with baseline trials. Both species reacted to positive (i.e., good food substituted for bad food) and negative (i.e., bad food substituted for good food) surprises. These results suggest that apes and dogs were able to individuate objects according to their properties or type in comparable ways. In addition, we looked for frustration and elation effects, but subjects' behaviors were not influenced by the food they saw and which they received in previous trials.

Keywords: object individuation, frustration effect, primates, domestic dogs

In 1928, Otto Leif Tinklepaugh raised the question of how nonhuman animals represent objects. In a series of experiments, he hid a piece of food under one of two containers while a macaque monkey watched. At the end of the trial, the monkey was invited to select a cup. Tinklepaugh varied when the cup was presented (different delays) and what was in the cup (the original food piece or a substitution). In one condition, the original food (banana) was substituted by another kind of food (lettuce). Tinklepaugh described the monkey's behavior as follows: "[the monkey] rushes to the proper container and picks it up. She extends her hand to seize the food. But her hand drops to the floor without touching it. She looks at the lettuce but (unless very hungry) does not touch it" (p. 224). On other occasions, when the subjects had seen the lettuce being hidden and they then found the lettuce, they consumed it without hesitation. Tinklepaugh concluded that the behavior of the tested monkeys (*Macaca fascicularis* and *Macaca mulatta*) was evidence of representative factors standing for the qualitative aspects of the food.

The question of how nonhuman primates represent and individuate objects was raised again 70 years later (Mendes, Rakoczy, &

Call, 2008; Phillips & Santos, 2007; Santos, Sulkowski, Spaepen, & Hauser, 2002; Uller, Carey, Hauser, & Xu, 1997). Mendes et al. (2008) studied object individuation in three great ape species with a refined manual search methodology. Similar to Tinklepaugh (1928), they tested whether apes individuate objects according to property or type of information: Subjects saw one kind of food reward being placed in a box and then either found that reward (expected) or a reward of a different kind (unexpected), which had been surreptitiously placed there by the experimenter. Subjects reached for the reward significantly more in the unexpected than in the expected condition, and they stopped searching when they found the original reward. The authors argued that, according to their findings, apes do not use simpler ways to track objects such as spatiotemporal tracking and feature placing but individuate objects according to their property or type. In other words, subjects did not search for just any reward but for exactly that reward that was placed in the box. Similar results were found using a looking time task with rhesus macaques (Phillips & Santos, 2007; Santos et al., 2002). Contrary to previous claims (Xu, 2002), it was concluded that individuating objects according to their properties or type was neither uniquely human nor essentially language dependent.

In all three recent studies, subjects were given different food with the same value. Tinklepaugh (1928), on the other hand, substituted less preferred food (lettuce) for preferred food (banana). In the majority of these cases, the monkeys kept searching for the banana and refused to eat the lettuce, although they accepted lettuce as a reward when this was the reward that had been hidden. This is an example of the much-documented *frustration effect* coming into play (Amsel & Roussel, 1952; see Papini & Dudley, 1997, and Flaherty, 1982, for reviews; Bentosela, Barrera,

Jakovcevic, Elgier, & Mustaca, 2008). When a familiar reward is no longer presented or is reduced in magnitude or quality, subjects often change their behavior. Crespi (1942) first reported that rats whose reward had been downgraded in one runway displayed an abrupt reduction in running speed to a level below the speed extrapolated from nonshifted control animals (Crespi, 1942; Flaherty, 1982). The same effect was found for licking behavior: If rats were given access to a tube containing 32% sucrose, and the concentration was then reduced to 4% sucrose, they licked less after the reduction than control rats that had access only to 4% sucrose (Vogel, Mikulka, & Spear, 1968). Capuchin monkeys also showed frustration in a similar situation: They rejected less preferred food if they had received more desirable food recently beforehand (Roma, Silberberg, Ruggiero, & Suomi, 2006), analogous to Tinklepaugh's monkeys that refused to eat the lettuce. But subjects might also change their behavior in another way when a familiar reward is no longer presented. Amsel and Roussel (1952) tested rats in two runways, each of them ending in a goal with a reward. They found that the rats ran faster in the second runway immediately after they were not rewarded in the first runway. They called this the frustration effect (Amsel & Roussel, 1952; Papini & Dudley, 1997), which might be analogous to Tinklepaugh's monkeys' searching behavior.

There is less evidence for animals changing their behavior when the reward is replaced by a preferred one. Tinklepaugh (1928) reported that in trials where the less preferred food (lettuce) was substituted with the preferred food (banana), his monkeys accepted the food immediately and showed no noticeable signs of "surprise." The rats tested by Crespi (1942), on the other hand, showed an increase in running speed—to a level above that of nonshifted, large-reward, control rats—when their reward was shifted from a smaller to a larger one. In his review, Flaherty (1982) came to the conclusion that this *elation effect* exists under certain circumstances, such as when there are enough acquisition trials and the reward is constantly delayed.

Despite considerable research focusing on dog cognition, little is known about how domestic dogs represent objects. Are dogs, like apes, capable of encoding the type of reward hidden in a particular location? Kaminski, Fischer, and Call (2008) found mixed results with two dogs that retrieved objects on command. Whereas one of the dogs appeared to encode the locations of multiple objects, the other did not. This is a particularly important question because although dogs show outstanding skills in the domain of communication with humans (Bräuer, Kaminski, Riedel, Call, & Tomasello, 2006; Hare, Brown, Williamson, & Tomasello, 2002; Kaminski, Tempelmann, Call, & Tomasello, 2009; Miklósi, Kubinyi, Gácsi, Virányi, & Csányi, 2003; Riedel, Schumann, Kaminski, Call, & Tomasello, 2008), they often perform poorly in studies involving causal relations and mental representations (Bräuer, Kaminski, et al., 2006; Erdőhegyi, Topál, Virányi, & Miklósi, 2007; Osthaus, Lea, & Slater, 2005; Rooijakkers, Kaminski, & Call, 2009).

The current study had two main aims. First, we investigated object individuation based on object properties in dogs and apes using an identical procedure. Would subjects show any measurable signs of "surprise" when finding a food item that did not correspond to the one they had seen the experimenter put into the container? Would they search for the piece of food that was hidden? The comparative nature of our study allowed us to assess whether object individuation was restricted to primates or could also be found in domestic dogs. Second, we examined frustration and elation effects, as we wanted to know whether the subjects' behavior was influenced by the food they had seen and that they received in the trials before. We predicted that subjects would reject the food when they were frustrated—when they had previously seen or received better food—and that they would eat the food sooner when they were elated, that is, when they had received less desirable food beforehand.

We used a container (magic cup) with a double bottom that allowed us to change the type of food that subjects discovered inside. We used a 2 × 2 design varying whether subjects received a generally preferred food and whether or not the food was substituted (surprise trials). For practical reasons, the studies of the apes and the dogs were conducted separately, with two different experimenters and in two different physical settings. The same methods, however, were used whenever possible, with deviations indicated where applicable.

Method

Subjects

Twenty great apes (eight chimpanzees [*Pan troglodytes*], five bonobos [*Pan paniscus*], four orangutans [*Pongo pygmaeus*], three gorillas [*Gorilla gorilla*]) of various ages (ranging from 6 to 34 years) were tested (see Table 1). All subjects lived in stable groups with their conspecifics in the Wolfgang Köhler Primate Research Center in the Leipzig Zoo. They had access to an outdoor area and an indoor area with natural vegetation and cages for the night. The apes were fed a diet of various fruits, vegetables, and cereals several times per day. Throughout testing, subjects were never food deprived and water was always available. Subjects could choose to stop participating at any time. They were tested individ-

Table 1

Name, Species, Gender, and Age of Great Apes Included in the Study

Subject	Species	Gender	Age (years)
Fraukje	Chimpanzee	F	31
Riet	Chimpanzee	F	30
Dorien	Chimpanzee	F	27
Frodo	Chimpanzee	M	14
Sandra	Chimpanzee	F[a]	14
Patrick	Chimpanzee	F	10
Pia	Chimpanzee	F	8
Lome	Chimpanzee	M	6
Joey	Bonobo	M	25
Ulindi	Bonobo	F	14
Limbuko	Bonobo	M	12
Kuno	Bonobo	M	11
Yasa	Bonobo	F	10
N'diki	Gorilla	F	30
Bebe	Gorilla	F	28
Viringika	Gorilla	F	12
Dunja	Orangutan	F	34
Pini	Orangutan	F	19
Dokana	Orangutan	F	18
Padana	Orangutan	F	10

[a] Neutered.

ually in familiar testing cages with one familiar experimenter. All apes had participated in a number of cognitive tests that were different from the current one.

Twenty domestic dogs (*Canis familiaris*), nine males and 11 females of various breeds and ages (ranging from 2 to 10 years) were tested (see Table 2). All subjects had lived their whole lives as pets with their owners. They had received only the normal obedience training typical of domestic dogs (except for two dogs that received agility and companion dog training). The dog owners were not present during testing and were not informed about the design of the experiments prior to testing. All dogs were tested individually by the same experimenter in the dog research facilities. Nearly all dogs had participated in different cognitive tests in the same facilities.

Materials

We used the so-called "magic cup" (see Figure 1), which was a double-walled metal container with a false bottom (10 cm × 5 cm) and a large lid. The cup was attached at its base to a wooden board (15 cm × 15 cm × 2 cm). During each trial, the human experimenter (E) placed the cup on a platform located between her and the subject. A large opaque board (40 cm × 60 cm) served to occlude the baiting process from the subject.

Apes were tested in a caged testing room (2.5 m × 2.2 m). The platform (85 cm × 32 cm) was fixed flush to the outside of the subjects' enclosure perpendicular to a Plexiglas testing panel. At the bottom of the Plexiglas panel were three holes positioned in a straight line. The magic cup stood in front of the middle hole. Grapes, apple pieces, and carrot pieces were used as rewards (each the size of a grape).

Dogs were tested in a dog testing room, sitting or standing opposite the experimenter. Small sausages, dry dog food, bread
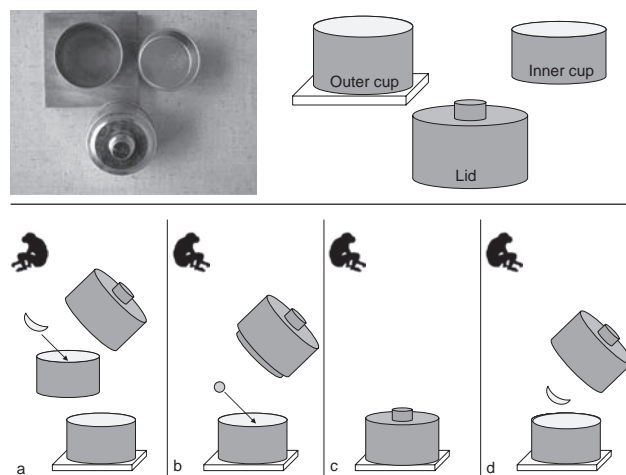
Table 2
Name, Breed, Gender, and Age of Dogs Included in the Study

Name	Breed	Gender	Age (years)
Amy	Magyar vizsla	F	3
Balou	Border collie × Labrador	M[a]	5
Bora	Labrador	F	6
Chico	Mixed breed	M	10
Emily	Labrador	F[a]	4
Higgins	Border collie × boxer	M[a]	8
Houk	Labrador × mixed breed	M	6
Jack	Jack Russell terrier	M	4
Karah	Labrador	F[a]	4
Koby	Labrador	M	6
Lea	Labrador	F[a]	2
Lucy	Labrador	F[a]	6
Luna	Dachshund × mixed breed	F[a]	2
Maica	Labrador	F	6
Marco	German shepherd × mixed breed	M[a]	8
Paula	Border collie × mixed breed	F	9
Rocky	American Staffordshire terrier × boxer	M[a]	6
Ronja	Sheltie × German shepherd	F	7
Tony	Bearded collie × mixed breed	F[a]	5
Yuri	Malinois × mixed breed	M[a]	4

[a] Neutered.

Figure 1. The magic cup and procedure: (a) baiting food in the inner cup out of view of the subject, (b) baiting food in the outer cup, (c) closing the lid, (d) opening lid so that only the inner cup is visible.

pieces, and carrot pieces were used as rewards (each the size of a large dog pellet).

Food Preference Test

The aim of this test was to find two different kinds of preferred food types for the subjects. The preferred food was always grapes for the apes and dog food sausages for the dogs. The inferior food needed to be less desirable than the preferred food, but subjects should still want to eat it. We asked the keepers and dog owners in advance about the subjects' food preferences, and we began the food preference tests with the least preferred food.

In the food preference test, E put a piece of highly desirable food (grape/sausage) on one side of the table and a piece of less desirable food (apples or carrots/dry dog food, bread, or carrots) on the other side. E then allowed the subjects to choose one. Apes indicated their choice by putting their finger through the hole in the Plexiglas panel, and dogs were allowed to take the food with their mouths. There were four trials of the preference test in which each kind of food was presented twice on each side. In the first two trials, subjects were only allowed to eat the food that they had selected, whereas in the last two trials, they could also eat the other piece of food.

Subjects were included in the study if they chose the preferred food 4 times in a row, as well as eating the less desirable food in the last two trials.

Procedure

E sat opposite the subject on a small chair. Prior to each trial, E baited a piece of food in the false bottom of the cup. This action was occluded by the board so that the subject could not see it. The trial started when E took a piece of food, held it in front of the subject, and put it in the cup, making sure the subject was watching as she did so. She then put the lid on the cup and opened it again after 1 s so that the subject now could see the first baited food rather than the second baited food that remained hidden in the false

bottom of the cup (see Figure 1). The subject was allowed to eat the food and its behavior was coded for 10 s.

Apes stood or sat behind the Plexiglas panel and were able to see the inside of the cup when they put their head close to the panel. As they were separated from E and could not reach the food by themselves, E put the food through the middle hole of the panel so that they could grab it with their hand or mouth. E tried to do this in the same manner and at the same speed in each trial, independent from subjects' behavior.

Dogs stood or sat opposite from E so that they could nearly reach the rim of the cup with their muzzle and were able to see the inside of the cup. To prevent dogs from eating the food, a second person held their collars held while E baited the cup and closed it. After E had opened the cup, the dog was held for an additional 1 s before being released to eat the food from the cup. This was to ensure that the latency to the moment subjects retrieved the food was comparable and so that dogs did not just grab the food immediately without looking at it. The second person could potentially lean forward and look inside the cup but tried to concentrate on the dog and tried to hold and release the dog in the same manner in each trial, independent from subjects' behavior.

Design

Subjects were presented with four different conditions depending on the type of food E baited in front of the subject (which would end up in the false bottom) and what kind of food was revealed when the lid was opened.

• BB—baseline: Bad food was baited in full view of the subject and bad food of the same shape and size appeared when the lid was opened.

• GG—baseline: Good food was baited in full view of the subject and good food of the same shape and size appeared when the lid was opened.

• GB—"negative surprise": Good food was baited in full view of the subject and bad food appeared when the lid was opened.

• BG—"positive surprise": Bad food was baited in full view of the subject and good food appeared when the lid was opened.

Although we were particularly interested in the "surprise" trials, we could not present them very often as subjects would quickly become habituated. For that reason, we administered a large number of baseline trials and interspersed them with some surprise trials. In each session, there were 16 trials: 14 baseline trials and two "surprise" trials (fourth or fifth trial and 12th or 13th trial—according to a predetermined coding sheet). After the eighth trial, there was a short break in each session.

Each subject was faced with four types of session depending on the kind of baseline and the type of surprise administered:

• BB-GB: baseline with bad food and "negative surprise,"
• BB-BG: baseline with bad food and "positive surprise,"
• GG-GB: baseline with good food and "negative surprise," and
• GG-BG: baseline with good food and "positive surprise."

Subjects were divided into four groups, each receiving their sessions in one of four orders (GG-BG/GG-GB/BB-BG/BB-GB or BB-GB/BB-BG/GG-GB/GG-BG or GG-GB/GG-BG/BB-GB/BB-BG or BB-BG/BB-GB/GG-BG/GG-GB). Apes received one session per day, whereas dogs received two with a break of at least 30 min in between. Thus, all together subjects were presented with four trials of each surprise condition (GB and BG) and 28 trials of each baseline condition (BB and GG). These were presented within 4 weeks. All trials were videorecorded.

Scoring, Reliability, and Analysis

We coded the behavior of the subjects for 10 s (or until it left the place in front of E in cases when it left), commencing the moment the lid was opened. We coded whether subjects (a) looked inside (apes) or smelled the cup (dogs), (b) begged by pointing toward the cup (apes), (c) left, and (d) when subjects started to eat. The behaviors were defined and scored as follows:

Checking the cup.
Looking inside. The number of trials the ape either had its head close to the window so that it had visual access to the inside of the cup or moved its head toward the window and gazed down to the cup.

Smelling. The number of trials in which the dog's nose was within 2 cm of the cup, the lid, or the table for at least 1 s, or in which the dog made sniffing noises. This behavior had to occur independent from eating the food.

Begging. The number of times per trial that the ape put its fingers or hand through one of the holes in the panel.

Leaving. The number of trials in which the subject left the area of approximately 1 m × 1 m in front of E where food could be acquired.

Latency to eat. The number of seconds that elapsed from the opening of the lid to the moment the subject started to eat, up to a maximum of 10 s—in cases when the subject did not start to eat during the timeframe that was coded. (Note that in the case of the apes, this depended on how quickly E handed over the reward. But as described earlier, E tried to hand over the reward in the same manner and within the same timeframe during each trial.)

Two independent observers, who were unaware of the purpose of the study, scored the behaviors of both species (one for the apes and one for the dogs), with the exception of the "looking inside" behavior, which was scored by the first author. Two other independent and naïve observers scored a randomly selected sample of 20% of the trials (one for the apes and one for the dogs) to assess interobserver reliability. Reliability was excellent (according to Fleiss, 1981) for apes (looking inside: Cohen's $\kappa = 0.76$, $n = 255$; begging: Pearson's $r = .71$, $n = 255$; leaving: Cohen's $\kappa = 0.90$, $n = 255$; latency to eat: Pearson's $r = .88$, $n = 255$) and dogs (smelling: Cohen's $\kappa = 0.80$, $n = 256$; leaving: Cohen's $\kappa = 0.87$, $n = 256$; latency to eat: Pearson's $r = .91$, $n = 256$).

We conducted the following analyses on each of the dependent measures for each species separately. First, we assessed the effect of the type of trial (surprise or baseline), the type of surprise that occurred during the session (positive or negative), and the type of baseline food (good or bad) using a repeated measures $2 \times 2 \times 2$ analysis of variance (ANOVA). Second, we analyzed whether the surprise trials had any influence on the subjects' baseline responses by comparing the behavior of the subjects in the baseline trials before and after the surprise trials in each session. As there were two surprise trials in each session, we concentrated on the initial eight trials prior to the break, which contained the first surprise trial (behavior before and after the surprise). We therefore used a repeated measures 2×4 ANOVA with the factors before–after (before vs. after surprise trial) and session type. Here, we also compared the

latency to eat the food as one of our questions was whether seeing the preferred food in the surprise trials would create an expectation of getting this food, therefore decreasing the acceptance of less desirable food.[1] We also analyzed whether there were gender differences and species differences between the apes as between-subjects factors. To keep the results clear, we report them only when they were significant.

We used repeated measures ANOVAs because the assumptions were met (independence of cases, interval data, normal distribution, and equality of variances). All comparisons are Bonferroni–Holm corrected.

Results

Type of Trial, Surprise, and Baseline Food

Apes. Figure 2 presents the mean percentage of trials in which the apes looked inside the cup as a function of the type of trial, surprise, and baseline food. The subjects looked inside the cup more in the surprise trials than in the baseline trials, $F(1, 19) = 26.49$, $\varepsilon^2 = 0.582$, $p < .001$, and when the surprise was negative, $F(1, 19) = 7.76$, $\varepsilon^2 = 0.29$, $p = .012$, but there was no effect for type of baseline food, $F(1, 19) = 0.01$, $\varepsilon^2 = 0.001$, $p = .913$. As there was a significant interaction for Type of Trial $\times$ Type of Surprise, $F(1, 19) = 6.06$, $\varepsilon^2 = 0.242$, $p = .024$, we conducted planned pairwise comparisons showing that apes looked inside the cup more often in the surprise than in the baseline trials, session GG-GB, $t(19) = 2.84$, $p = .020$; session BB-BG, $t(19) = 3.38$, $p = .009$; session BB-BG, $t(19) = 4.38$, $p < .001$, but not when a positive surprise was embedded in the baseline trials with good food, session GG-BG, $t(19) = 0.93$, $p = .365$. All other interactions were not significant ($ps > .05$).

Figure 3 presents the mean percentage of trials in which the apes left their place in front of E as a function of the type of trial, surprise, and baseline food. Apes left more often when there was bad food in the baseline trials, $F(1, 19) = 5.11$, $\varepsilon^2 = 0.212$, $p = .036$. They also left more often in baseline trials than in surprise trials, $F(1, 19) = 7.78$, $\varepsilon^2 = 0.290$, $p = .012$, but there was no effect for the type of surprise, $F(1, 19) = 0.06$, $\varepsilon^2 = 0.003$, $p =$
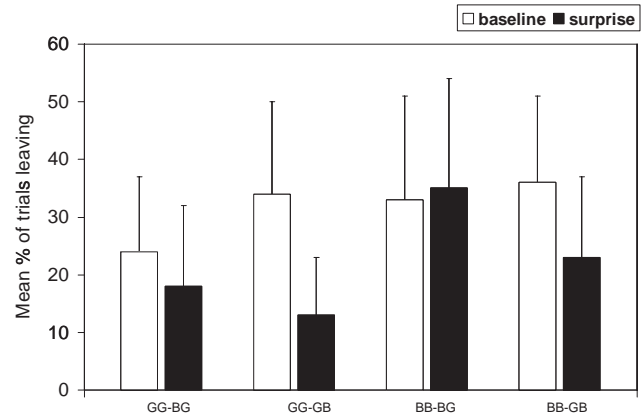
Figure 3. Mean percentage of trials (+95% confidence interval) in which apes left their place in front of the experimenter, shown separately for surprise trials and baseline trials (G = good food; B = bad food).

.816. However, there was an interaction effect for Type of Trial $\times$ Type of Surprise, $F(1, 19) = 6.09$, $\varepsilon^2 = 0.243$, $p = .023$. Planned pairwise comparisons showed that, within the surprise trials, apes remained longer than in baseline trials when there was a bad surprise embedded in a session with good food, session GG-GB, $t(19) = -3.69$, $p = .008$, but not in the other sessions, session GG-BG, $t(19) = -1.62$, $p = .244$; session BB-BG, $t(19) = -2.06$, $p = .159$; session BB-BG, $t(19) = 0.35$, $p = .728$. Moreover, male apes left the place in front of E more often than females, $F(1, 18) = 5.17$, $\varepsilon^2 = 0.223$, $p = .035$.

As apes "looked inside" more often and left the scenario less often when there was a surprise, it is possible that these two measures depended on one another. Thus, it is possible that we found more frequent looking simply because the apes were present more often. For this reason, we reanalyzed looking behavior excluding all trials in which the subjects left their position in front of E. We found similar results: Subjects looked inside the cup more in surprise trials than in baseline trials, $F(1, 19) = 8.15$, $\varepsilon^2 = 0.300$, $p = .010$, and if the surprise was negative, $F(1, 19) = 10.05$, $\varepsilon^2 = 0.346$, $p = .005$, but there was no effect for type of baseline, $F(1, 19) = 0.463$, $\varepsilon^2 = 0.024$, $p = .504$. As the interaction between type of trial and type of surprise was significant, $F(1, 19) = 7.27$, $\varepsilon^2 = 0.277$, $p = .014$, we conducted planned pairwise comparisons, which showed that apes looked inside the cup more often than in baseline trials when the surprise was negative, session GG-GB, $t(19) = 3.31$, $p = .016$; session BB-BG, $t(19) = 2.80$, $p = .036$. However, they did not do so when the surprise was positive, GG-BG, $t(19) = 0.49$, $p = .628$; session BB-BG, $t(19) = -0.79$, $p = .884$. All other interactions were not significant. Apes also begged more in surprise trials than in baseline trials, $F(1, 19) = 34.53$, $\varepsilon^2 = 0.645$, $p < .001$. No other comparisons were significant ($ps > .05$).

Dogs. Figure 4 presents the mean percentage of trials in which dogs smelled near the table, the cup, or the lid as a function
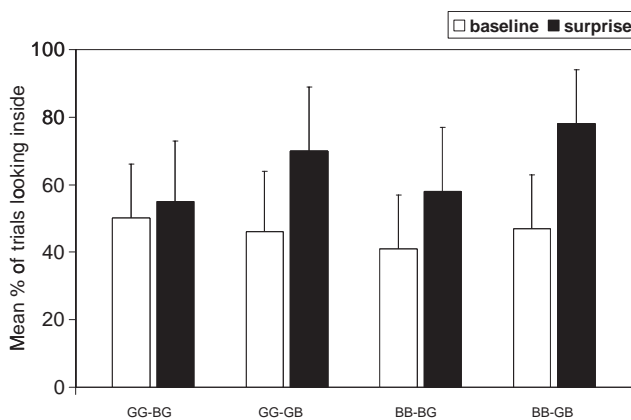

Figure 2. Mean percentage of trials (+95% confidence interval) in which apes looked inside the cup, shown separately for surprise trials and baseline trials (G = good food; B = bad food).

---

[1] To keep the results clear, we present the measure "latency to eat" only in the "behavior before and after the surprise" analysis. We did all analyses with that measure, but we found no significant results except for one interaction.
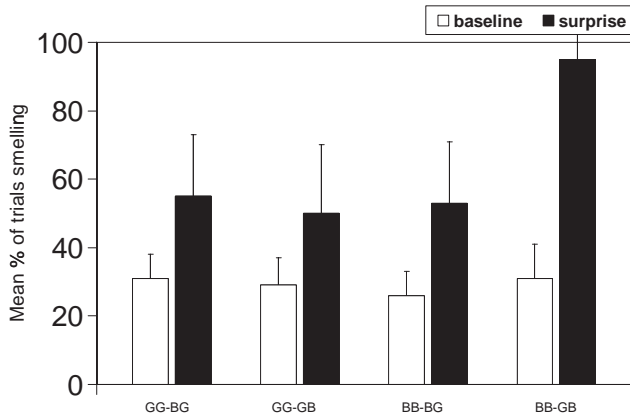
Figure 4. Mean percentage of trials (+95% confidence interval) in which dogs smelled the cup, the lid, or the table, shown separately for surprise trials and baseline trials (G = good food; B = bad food).

of the type of trial, surprise, and baseline food. Subjects smelled more often in surprise trials than in baseline trials, $F(1, 19) = 60.64$, $\varepsilon^2 = 0.760$, $p < .001$, and if the surprise was negative, $F(1, 19) = 5.11$, $\varepsilon^2 = 0.212$, $p = .036$, but there was no effect for type of baseline, $F(1, 19) = 4.02$, $\varepsilon^2 = 0.175$, $p = .059$. There were interaction effects for Type of Baseline × Type of Surprise, $F(1, 19) = 13.25$, $\varepsilon^2 = 0.411$, $p = .002$, for Type of Baseline × Type of Trial, $F(1, 19) = 9.56$, $\varepsilon^2 = 0.335$, $p = .006$, and for Type of Baseline × Type of Surprise × Type of Trial, $F(1, 19) = 5.95$, $\varepsilon^2 = 0.238$, $p = .025$. Planned pairwise comparisons showed that dogs smelled more in surprise trials than in baseline trials if they received less desirable food in the baseline trial, session BB-BG, $t(19) = -3.53$, $p = .012$; session BB-GB, $t(19) = -9.72$, $p < .001$, but not if they received desirable food in the baseline trial, session GG-GB, $t(19) = 2.07$, $p = .212$; session GG-BG, $t(19) = 2.58$, $p = .09$. They smelled more often in a negative surprise trial when it was embedded in baseline trials with bad food than in baseline trials with good food, session BBGB versus GGGB, $t(19) = -4.41$, $p < .001$. Moreover, in baseline trials with bad food, they smelled more after a negative surprise than after a positive surprise, session BBGB versus BBBG, $t(19) = -4.68$, $p < .001$. No other comparisons were significant ($ps > .05$).

Figure 5 presents the mean percentage of trials in which dogs left their place in front of E as a function of the type of trial, surprise, and baseline food. Dogs left more often in the baseline than surprise trials; there was an effect for the type of trial, $F(1, 19) = 9.14$, $\varepsilon^2 = 0.325$, $p = .007$, but no effect for the type of surprise, $F(1, 19) = 2.87$, $\varepsilon^2 = 0.131$, $p = .107$, and type of baseline, $F(1, 19) = 0.39$, $\varepsilon^2 = 0.020$, $p = .538$. However, there was an interaction effect between the type of trial and type of surprise, $F(1, 19) = 12.65$, $\varepsilon^2 = 0.400$, $p = .002$. Planned pairwise comparisons showed that dogs left more only in baseline than surprise trials when the surprise was negative, session GG-GB, $t(19) = -2.82$, $p = .033$; session BB-GB, $t(19) = 3.41$, $p = .012$, but not when there was a positive surprise, session GG-BG, $t(19) = -0.76$, $p = .916$; session BB-BG, $t(19) = 0.56$, $p = .583$. All other interactions were not significant ($ps > .05$). Moreover, male dogs left the place in front of E more often than females, $F(1, 18) = 7.04$, $\varepsilon^2 = 0.281$, $p = .016$.

As dogs looked inside the cup more often and left their position less often when there was a surprise, it is possible that these two measures depended on one another. Individual dogs may appear to be smelling more frequently simply because they were present longer. For this reason, we reanalyzed smelling behavior by excluding all trials in which the subjects left their place in front of E. We found similar results: Subjects smelled more often in the surprise trials than in the baseline trials, $F(1, 19) = 38.86$, $\varepsilon^2 = 0.672$, $p < .001$, and when the food in the baseline was bad, $F(1, 19) = 8.58$, $\varepsilon^2 = 0.311$, $p = .009$, but there was no effect for the type of surprise, $F(1, 19) = 2.00$, $\varepsilon^2 = 0.095$, $p = .174$. There were interaction effects between type of baseline and type of surprise, $F(1, 19) = 11.64$, $\varepsilon^2 = 0.380$, $p = .003$, Type of Baseline × Type of Trial, $F(1, 19) = 11.57$, $\varepsilon^2 = 0.378$, $p = .003$, and Type of Baseline × Type of Surprise × Type of Trial, $F(1, 19) = 5.98$, $\varepsilon^2 = 0.238$, $p = .025$. Planned pairwise comparisons showed that dogs smelled more often in surprise trials than in baseline trials when they got bad food in that session, session BB-BG, $t(19) = -3.70$, $p = .012$; session BB-GB, $t(19) = -9.28$, $p < .001$, but not when they got good food, session GG-GB, $t(19) = 1.64$, $p = .351$; session GG-BG, $t(19) = 2.38$, $p = .112$. They smelled more often in a negative surprise trial when it was embedded in baseline trials with bad food than in baseline trials with good food, session BBGB versus GGGB, $t(19) = -4.60$, $p < .001$. Moreover, in baseline trials with bad food, they smelled more after a negative surprise than after a positive surprise, session BBGB versus BBBG, $t(19) = -3.68$, $p = .01$. No other comparisons were significant ($ps > .05$).

Behavior Before and After the Surprise in Baseline Trials

Apes. Apes looked inside the cup more during baseline trials before the surprise trials than after them. There was an effect for before–after, $F(1, 19) = 5.29$, $\varepsilon^2 = 0.218$, $p = .033$, but not for session type, $F(1, 19) = 2.58$, $\varepsilon^2 = 0.120$, $p = .062$, and we found no Before–After × Session Type effect, $F(1, 19) = 0.92$, $\varepsilon^2 = 0.046$, $p = .438$. In contrast, there was no difference in the percentage of trials in which apes left their place in front of E.
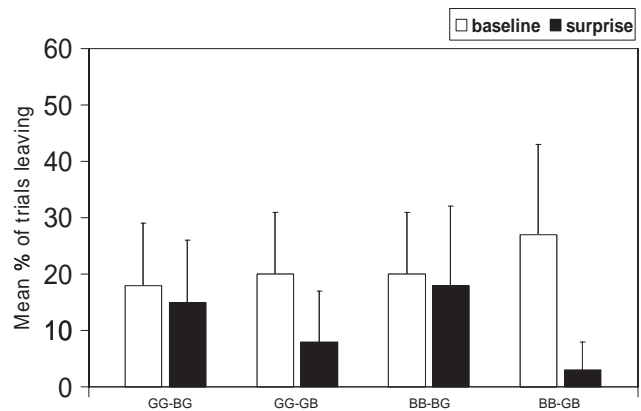

Figure 5. Mean percentage of trials (+95% confidence interval) in which dogs left their place in front of the experimenter, shown separately for trials and baseline trials (G = good food; B = bad food).

There was no effect for session type, $F(1, 19) = 0.86$, $\varepsilon^2 = 0.043$, $p = .470$, for the factor before–after, $F(1, 19) = 4.04$, $\varepsilon^2 = 0.175$, $p = .059$, and for Before–After × Session Type, $F(1, 19) = 1.65$, $\varepsilon^2 = 0.080$, $p = .188$. Regarding begging behavior, the repeated measures ANOVA also showed no effect: before–after, $F(1, 19) = 0.23$, $\varepsilon^2 = 0.012$, $p = .641$; session type, $F(1, 19) = 0.32$, $\varepsilon^2 = 0.016$, $p = .813$; Before–After × Session Type, $F(1, 19) = 0.88$, $\varepsilon^2 = 0.044$, $p = .455$.

We also compared the latency to eat the food to find out whether there was a change in subjects' acceptance of the bad food. However, apes showed no difference in their latency to eat the food. We found no effect for the factors before–after, $F(1, 19) = 3.18$, $\varepsilon^2 = 0.143$, $p = .090$, session type, $F(1, 19) = 2.27$, $\varepsilon^2 = 0.107$, $p = .090$, and Before–After × Session Type, $F(1, 19) = 2.09$, $\varepsilon^2 = 0.099$, $p = .111$.

Dogs. Subjects displayed no differences in smelling behavior during baseline trials before and after the surprise trials. There was no effect for before–after, $F(1, 19) = 3.83$, $\varepsilon^2 = 0.168$, $p = .065$, session type, $F(1, 19) = 0.64$, $\varepsilon^2 = 0.033$, $p = .590$, or Before–After × Session Type, $F(1, 19) = 0.70$, $\varepsilon^2 = 0.034$, $p = .574$.

Regarding the trials in which dogs left their place in front of E, there was no effect for before–after, $F(1, 19) = 0.29$, $\varepsilon^2 = 0.015$, $p = .597$, and session type, $F(1, 19) = 1.45$, $\varepsilon^2 = 0.071$, $p = .239$, although there was a significant interaction, $F(1, 19) = 3.57$, $\varepsilon^2 = 0.158$, $p = .019$. However, planned pairwise comparisons revealed no differences between the baseline trials before and after surprise trials in any session: session GG-BG, $t(19) = 1.35$, $p = .384$; session GG-GB, $t(19) = -0.47$, $p = .641$; session BB-BG, $t(19) = 1.81$, $p = .344$; session BB-GB, $t(19) = -1.51$, $p = .444$. We also found no effect for latency to eat for before–after, $F(1, 19) = 0.22$, $\varepsilon^2 = 0.011$, $p = .646$; session type, $F(1, 19) = 2.04$, $\varepsilon^2 = 0.097$, $p = .118$; or Before–After × Session Type, $F(1, 19) = 0.13$, $\varepsilon^2 = 0.007$, $p = .943$.

Discussion

Apes looked inside the cup more often and dogs smelled inside and around the cup more frequently when one kind of food was substituted for another kind of food. Both species also remained close to the experimenter more often, and apes begged more in trials when there had been a surprise. Apes and dogs reacted to both positive (i.e., good food substituted for bad food) and negative (i.e., bad food substituted for good food) surprises. Nevertheless, both apes and dogs reacted more strongly to negative rather than positive surprises. Taken together, these results suggest that apes and dogs were able to individuate objects according to their properties or type. In other words, subjects not only remembered that a piece of food was placed in the magic cup, but they expected to find exactly that piece of food when the cup was opened again.

One remarkable finding of this study is that apes and dogs showed very similar behavioral patterns: They looked or smelled more and stayed close to the experimenter more often when there was a surprise. Thus, the differences that have been found between apes and dogs in studies of invisible object displacements (Collier-Baker, Davis, & Suddendorf, 2004; Fiset & LeBlanc, 2007; Rooijakkers et al., 2009; Watson et al., 2001) cannot be attributed to differences in their ability to individuate objects. It is conceivable therefore that object individuation based on kind, which emerges

in human ontogeny within the first year, might be a basic skill that exists in many animal species.

Subjects reacted to both positive and negative surprises; however, they reacted more often to the negative than the positive surprises. This contrasts with Tinklepaugh (1928), who found an effect only for negative surprises. However, a surprising result in the current study was that there was no difference at all in the latency from the moment the lid was opened to the subject starting to eat. Our predictions were that subjects would hesitate to eat the food or reject the food when they were frustrated, that is, when they had previously seen or received better food; and that subjects would eat the food sooner when they were elated, that is, when they had received less desirable food beforehand. However, unlike Tinklepaugh's monkeys, apes and dogs never rejected the less desirable food. They also never hesitated to eat when there was a negative surprise. In addition, there was also no change in their acceptance of the food before and after the surprise trial.

There are two possible reasons why we could not detect the frustration and elation effects. One concerns the number of trials. In studies of frustration and elation effects, subjects usually receive a large number of trials in which they learn to receive a certain reward. Tinklepaugh's (1928) monkeys, for example, received hundreds of trials with banana pieces before the food was substituted. In the current experiment, subjects received one reward type over just three or four trials before seeing or receiving the other type of reward in a surprise trial. In other words, subjects were not given the opportunity to learn to always expect one type of reward and were thus not frustrated or elated when they saw or received the substitute.

The second possible reason why we did not find frustration and elation may lie in the pretest. The aim of the pretest was to find two different kinds of preferred food types for the subjects while making sure that subjects would still eat the less desirable food. Subjects were included in the study only when they had proven that they would eat the less desirable food. Thus, although subjects clearly preferred the good food over the bad food, their acceptance of the less preferable food was perhaps still too high, causing them never to reject it. Studies with apes on inequity aversion have produced very different refusal rates for less preferred food such as carrots and apples (Bräuer, Call, & Tomasello, 2006, 2009; Brosnan, Schiff, & de Waal, 2005). It should also be noted that the same apes used in this study had previously rejected the less desirable food in the studies of Bräuer, Call, et al. (2006, 2009), despite similar preconditions for participation. However, the apes in that study had to accept the less desirable food in the presence of a highly preferable food. Thus, in a different competitive situation with a conspecific, our apes do sometimes reject the same less desirable food. Note also that Tinklepaugh's monkeys usually accepted lettuce when this was the hidden reward. The high acceptance of the less preferred reward cannot, therefore, be the only reason that we did not find instances of frustration, at least not for the apes. But it is conceivable that both reasons previously discussed could account for the lack of frustration and elation effects; namely, the small number of trials and the high acceptance of less desirable food.

We found that both species stayed close to the experimenter more in the surprise trials compared with baseline trials, and they displayed increased begging (apes), looking (apes), and smelling (dogs) behaviors. It is likely that this reflects the subjects' expec-

tations that the obscured food was still there and that they would receive it. This is not only supported by previous studies on object individuation in primates where subjects searched for the missing food (Mendes et al., 2008; Phillips & Santos, 2007; Santos et al., 2002), but also by other findings. For example, Watanabe et al. (2001) interpreted behavioral reactions such as anticipatory licks as reward expectations in monkeys. Bräuer, Call, & Tomasello (2006) also found that apes remained longer when a conspecific received better food than themselves. The authors interpreted this as an indication that seeing another individual receive high-quality food creates the expectation of receiving the same food oneself. Similarly, Call, Hare, Carpenter, and Tomasello (2004) found that chimpanzees were more likely to stay with an experimenter who was unable, as opposed to unwilling, to give them food. Subjects in the current study also seemed to have the expectation that they would receive the food piece that had been obscured.

One could, however, argue that in the dogs' case they put their nose close to the cup, the lid, and the table more often when desirable food was substituted with less desirable food, not because they expected to find the desirable food that had disappeared but because they could still smell it hidden in the compartment of the magic cup. But we can exclude this possibility because dogs showed no increase in smelling behavior in the baseline with good food, in which good food remained hidden in the cup after they had eaten the reward. One could still argue that the dogs just reacted to a change of the smell. But then we would expect increased smelling behavior in both sessions in which dogs had smelled one kind of food in all baseline trials and after the surprise trial there was another "new" food in the cup (sessions GG-BG and BB-GB). However, this was not the case. It is still remarkable that dogs reacted most after the negative surprise when the baseline was bad food (session BB-GB). We can only speculate to explain that finding but it might be the case that dogs reacted more as seeing the good food—and not getting it—might be much more frustrating when they had received only bad food during the whole session compared with a session in which they got good food and only one of these disappeared (session GG-GB).

In conclusion, we have demonstrated that apes and dogs are capable of individuating objects according to their properties or type, showing similar behavioral patterns. More species will need to be tested to find out whether this is a basic mammalian skill and to learn about the mechanisms involved. We were unable to detect frustration and elation effects because subjects neither hesitated before eating nor were faster to begin eating when they had seen or received a different type of food beforehand.

References

Amsel, A., & Roussel, J. (1952). Motivational properties of frustration: I. Effect on a running response of the addition of frustration to the motivational complex. *Journal of Experimental Psychology: General, 43,* 363–368.

Bentosela, M., Barrera, G., Jakovcevic, A., Elgier, A. M., & Mustaca, A. E. (2008). Effect of reinforcement, reinforcer omission and extinction on a communicative response in domestic dogs (*Canis familiaris*). *Behavioural Processes, 78,* 464–469.

Bräuer, J., Call, J., & Tomasello, M. (2006). Are apes really inequity averse? *Proceedings of the Royal Society B: Biological Sciences, 273,* 3123–3128.

Bräuer, J., Call, J., & Tomasello, M. (2009). Are apes inequity averse?

New data on the token-exchange paradigm. *American Journal of Primatology, 71,* 175–181.

Bräuer, J., Kaminski, J., Riedel, 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.

Brosnan, S. F., Schiff, H. C., & de Waal, F. B. M. (2005). Tolerance for inequity may increase with social closeness in chimpanzees. *Proceedings of the Royal Society B: Biological Sciences, 272,* 253–258.

Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). "Unwilling" versus "unable": Chimpanzees' understanding of human intentional action. *Developmental Science, 7,* 488–498.

Collier-Baker, E., Davis, J. M., & Suddendorf, T. (2004). Do dogs (*Canis familiaris*) understand invisible displacement? *Journal of Comparative Psychology, 118,* 421–433.

Crespi, L. P. (1942). Quantitative variation in incentive and performance in the white rat. *American Journal of Primatology, 55,* 467–517.

Erdőhegyi, Á., Topál, J., Virányi, Z., & Miklósi, Á. (2007). Dog-logic: Inferential reasoning in a two-way choice task and its restricted use. *Animal Behaviour, 74,* 725–737.

Fiset, S., & LeBlanc, V. (2007). Invisible displacement understanding in domestic dogs (*Canis familiaris*): The role of visual cues in search behavior. *Animal Cognition, 10,* 211–224.

Flaherty, C. E. (1982). Incentive contrast: A review of behavioral changes following shifts in reward. *Animal Learning & Behavior, 10,* 409–440.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions.* New York: Wiley.

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002, November 22). The domestication of social cognition in dogs. *Science, 298,* 1634–1636.

Kaminski, J., Fischer, J., & Call, J. (2008). Prospective object search in dogs: Mixed evidence for knowledge of what and where. *Animal Cognition, 11,* 367–371.

Kaminski, J., Tempelmann, S., Call, J., & Tomasello, M. (2009). Domestic dogs comprehend human communication with iconic signs. *Developmental Science, 12,* 831–837.

Mendes, N., Rakoczy, H., & Call, J. (2008). Ape metaphysics: Object individuation without language. *Cognition, 106,* 730–749.

Miklósi, Á., Kubinyi, E., Gácsi, M., Virányi, Z., & Csányi, V. (2003). A simple reason for a big difference: Wolves do not look back at humans but dogs do. *Current Biology, 13,* 763–766.

Osthaus, B., Lea, S. E. G., & Slater, A. M. (2005). Dogs (*Canis lupus familiaris*) fail to show understanding of means–end connections in a string-pulling task. *Animal Cognition, 8,* 37–47.

Papini, M. R., & Dudley, R. T. (1997). Consequences of surprising reward omissions. *Review of General Psychology, 1,* 175–197.

Phillips, W., & Santos, L. R. (2007). Evidence for kind representations in the absence of language: Experiments with rhesus monkeys (*Macaca mulatta*). *Cognition, 102,* 455–463.

Riedel, J., Schumann, K., Kaminski, J., Call, J., & Tomasello, M. (2008). The early ontogeny of human–dog communication. *Animal Behaviour, 75,* 1003–1014.

Roma, P. G., Silberberg, A., Ruggiero, A. M., & Suomi, S. J. (2006). Capuchin monkeys, inequity aversion, and the frustration effect. *Journal of Comparative Psychology, 120,* 67–73.

Rooijakkers, E., Kaminski, J., & Call, J. (2009). Comparing dogs and great apes in their ability to visually track object transpositions. *Animal Cognition, 12,* 789–796.

Santos, L. R., Sulkowski, G. M., Spaepen, G. M., & Hauser, M. D. (2002). Object individuation using property/kind information in rhesus macaques (*Macaca mulatta*). *Cognition, 83,* 241–264.

Tinklepaugh, O. L. (1928). An experimental study of representative factors in monkeys. *Journal of Comparative Psychology, 8,* 197–236.

Uller, C., Carey, S., Hauser, M., & Xu, F. (1997). *Is language needed for constructing sortal concepts? A study with non-human primates*. Sommerville, MA: Cascadilla Press.

Vogel, J. R., Mikulka, P. J., & Spear, N. E. (1968). Effects of shifts in sucrose and saccharin concentrations on licking behavior in the rat. *Journal of Comparative and Physiological Psychology, 66,* 661–666.

Watanabe, M., Cromwell, H. C., Trembley, L., Hollermann, J. R., Hikosaka, K., & Schultz, W. (2001). Behavioral reactions reflecting differential reward expectations in monkeys. *Experimental Brain Research, 140,* 511–518.

Watson, J. S., Gergely, G., Csányi, V., Topál, J., Gácsi, M., & Sárközi, Z. (2001). Distinguishing logic from association in the solution of an invisible displacement task by children (*Homo sapiens*) and dogs (*Canis familiaris*): Using negation of disjunction. *Journal of Comparative Psychology, 115,* 219–226.

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy cognition. *Cognition, 85,* 223–250.