Running Head: ARE APES ESSENTIALISTS

Are apes essentialists?

Scope and limits of psychological essentialism in great apes

Trix Cacchione[1], Christine Hrubesch[1,2,3], Josep Call[2,4], & Hannes Rakoczy[5]

[1] Department of Psychology, University of Berne, Switzerland

[2] Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

[3] Anthropological Institute & Museum, University of Zurich, Switzerland

[4] School of Psychology & Neuroscience, University of St Andrews, United Kingdom

[5] Department of Developmental Psychology, University of Goettingen, Germany

Correspondence concerning this article should be addressed to Trix Cacchione at the Department of Psychology, University of Berne, Fabrikstrasse 8, CH-3012 Bern; Phone: ++41 (0) 31 631 58 96; E-Mail: trix.cacchione@psy.unibe.ch

Abstract

Human reasoning is characterized by psychological essentialism (Gelman, 2003): when reasoning about objects we distinguish between deep essential properties defining the object’s kind and identity, and merely superficial features that can be changed without altering the object’s identity. To date, it is unclear whether psychological essentialism is based on the acquisition of linguistic means (such as kind terms) and therefore uniquely human, or whether it is a more fundamental cognitive capacity which might be present also in the absence of language. In the present study, we addressed this question by testing whether, and if so, under which circumstances non-human apes also rely on psychological essentialism to identify objects. For this purpose, we adapted classical verbal transformation scenarios used in research on psychological essentialism (Keil, 1989), and implemented them in two non-verbal tasks: first, a box task, typically used to test object individuation (Experiment 1), and second, an object choice task, typically used to test object discrimination, object preferences and logical inferences (Experiments 2-4). Taken together, the results of the four experiments suggest that under suitable circumstances (when memory and other task demands are minimized), great apes engage in basic forms of essentialist reasoning. Psychological essentialism is thus possible also in the absence of language.

Key words: Object individuation, Essentialism, Comparative cognition, Categorization, Conceptual development

1. Introduction

Adult human thinking is characterized by psychological essentialism (Gelman, 2003; Keil, 1989): people think about objects and kinds in the way philosophers have argued the semantics of natural kinds works (Kripke, 1972; Putnam, 1975). Humans conceive of natural kinds (e.g., chemical kinds like “diamond”, or biological kinds like “tiger”) as having two levels of properties: deep essential properties that define the objects’ kind and identity, and merely superficial features that can be changed without altering the objects’ identity. Deep essential properties vary from domain to domain – they might be deep chemical properties in the case of diamonds, for example, or deep biological properties in the case of tigers. Crucially, however, humans believe that kinds (e.g., tigers, diamonds, roses) possess essential properties, without actually knowing what these essences are – instead, they mostly operate with “essence placeholder” conceptions (see e.g., Medin & Ortony, 1989). And though humans are largely ignorant about what exactly the essential properties of a given natural kind might be, one central folk assumption is that essential properties tend to lie deep within objects rather than being visible from the surface. A natural kinds’ identity, therefore, is thought to be changed when its inside is changed, but not when its surface appearance is changed (Gelman & Wellman, 1991; Gottfried & Gelman, 2005; Newman, Herrmann, Wynn & Keil, 2008).

Much experimental work has shown that from around four years of age human reasoning is consistent with psychological essentialism (see Gelman, 2003; Keil, 1982, for a review). First, in category-based induction tasks, children and adults readily transfer non-visible internal features and functions (e.g., possessing a certain organ) from one exemplar of a given kind to new members of the same category. Importantly, this happens also if category-membership competes with perceptual similarity (e.g., if a newly encountered exemplar shares more surface similarity with non-members than with other members of the same category; Gelman & Markman, 1986, 1987). For instance, 4-year-old children infer that sharks breathe as tropical fish rather than dolphins, because sharks are fish despite looking more similar to dolphins (Gelman & Markman, 1986). Second, children and adults judge the trans-temporal identity of objects of a certain kind based on the continuity of their essential properties, neglecting transformations of superficial features. Subjects presented with a token of a natural kind (e.g., a raccoon), for example, judge that superficial transformations (e.g., being painted like a squirrel, growing up among squirrels, learning to make squirrel sounds) do not alter the identity of the token, which remains a raccoon even if it looks like a squirrel: Keil, 1989).

All of this work, however, is heavily dependent on linguistic material and measures. Little is known, therefore, about potential pre-linguistic cognitive roots of psychological essentialism, both ontogenetically and phylogenetically. One interesting possibility that motivates the present paper is that certain forms of keeping track and re-identifying objects over time, and despite changes in superficial appearance, may entail a basic and primordial form of essentialist reasoning: such re-identification already involves a distinction between persisting essential properties that secure identity and changing surface features. The clearest case of such object cognition is the so-called sortal individuation of objects, in particular of natural kind objects. Sortal natural kind concepts such as DOG, APPLE or ELM TREE, lexicalized as count nouns in classifier languages ("a dog" etc.), supply criteria for individuation ("How many dogs are in this room?") and identification ("Is this the same dog as the one seen before?") (Xu, 2007). Without doubt, adult sortal individuation of natural kind objects, as measured verbally in classical transformation vignettes (Keil, 1989), clearly embodies essentialist reasoning ("this is still the same racoon as the one seen before even though it looks completely different now").

However, developmental and comparative work has also investigated sortal object individuation in preverbal humans and non-verbal primates with non-linguistic methods. In his classic studies, for instance, Herrnstein presented pigeons with pairs of pictures and trained them to respond to the presence or absence of specific stimuli in the pictures (e.g., humans, trees; Herrnstein & Loveland, 1964; Herrnstein, Loveland, & Cable, 1976). When presented with new pairs of pictures, pigeons could successfully discriminate the ones containing the target stimulus, although this never looked exactly the same, suggesting that pigeons processed kind/category of the displayed entities rather than mere surface features. Also African grey parrots apparently process categories when counting the number of similar (i.e., belonging to the same-category) items in a presentation, despite their differences in surface features (see Pepperberg, 2013).

In a typical developmental paradigm, infants or non-human primates saw an object of Kind 1 disappear in a box, and were then allowed to search for the object in the box. Depending on the condition, infants/non-human primates found either an object of Kind 1 (same kind condition) or an object of Kind 2 (different kind condition). Infants from around 12 months of age (van de Walle, Carey, & Prevor, 2000; Krøjgaard, 2004; Xu & Baker, 2005; Xu, Carey, & Quint, 2004) and non-human primates (Mendes, Rakoczy, & Call, 2008, 2011; Philipps & Santos, 2007; Santos, Sulkowski, Spaepen, & Hauser, 2002) searched significantly longer in the box in the different kind condition, as compared to the same kind condition. These findings, however, are open to different interpretations. First, they might suggest that subjects individuate objects in terms of their natural kinds, i.e., in terms of their deep essential properties that determine their trans-temporal identity and survive changes in merely superficial properties. However, a more parsimonious explanation is also possible. Given that in normal circumstances the essential properties (i.e., what kind an object belongs to) and the superficial features (i.e., what it looks like) are confounded, subjects could have simply based their numerical expectations on the representation of superficial features, perhaps by using prototype or other feature-based representations of the object categories in question.

Additional studies on human infants, however, speak against such an explanation. In an object individuation task, Cacchione, Schaub, and Rakoczy (2013) found evidence that 14-month-old infants distinguished between superficial feature transformations that were diagnostic of identity changes and those that were not. In their study, they adopted the box paradigm described above but used a special type of toys. They could be opened by a zip fastener and turned inside out, e.g., looking either like a pig or like a ball. Before the experiment, half of the infants had been familiarised with one the toys and therefore knew that the toys could be turned inside out (and that e.g. the pig and the ball were really one and the same object). The other half of infants was unaware of that. In the test, the experimenter hid another one of the toys in the box, moved its hands as if turning it and let the infants search the box. They either found the toy as having the same features (i.e., they saw the toy entering like a carrot and then found a carrot), or they found the toy as having different features (i.e., they saw the toy entering like a carrot and then found a bunny). While the unfamiliarised infants considered the feature differences to be diagnostic for an identity change (i.e., they searched in the box for another object when the toy they retrieved looked different from the one entering the box), the familiarised infants did not. Therefore, infants did not merely track visual features in this task. Instead they interpreted the observed feature differences as related to hidden causal/functional attributes of a specific kind of object that they encountered in the prior training. Converging evidence for a systematic distinction between deep and superficial features at this age comes from a study by Newman and colleagues (2008). In this study, 14-month-old infants associated the behaviour of a toy cat with its internal (deep) rather than with its external (superficial) features, appreciating that the behaviour of an agent is more likely to be caused by internal properties rather than by more accidental external features.

From a comparative point of view, two recent studies have attempted to tease apart the representation of superficial and essential features (Phillips & Santos, 2007; Phillips, Shankar & Santos, 2010). In the first study, monkeys observed an experimenter pretending to cut a small piece from a familiar food (e.g., a coconut or an apple) and placing it inside a box (Phillips & Santos, 2007). In fact, however, what the monkeys really saw being placed into the box by the experimenter was a neutral white piece of plastic that visually looked equally like a piece of coconut as like a piece of apple. When searching in the box, monkeys either found a consistent piece of fruit (e.g., coconut after having seen the experimenter cut from a coconut) or an inconsistent one (e.g., apple after the same demonstration). Monkeys searched longer after finding the inconsistent kind of fruit, suggesting that they represented the object based on its internal properties (and thus searched longer when unexpectedly finding an object with different internal properties), rather than based on its immediate perceptual properties (which were identical both in consistent and inconsistent test conditions). In the second study, monkeys first saw an object of Kind 1 (e.g., an apple) and then saw that object disappear behind another object of Kind 2 (e.g., a coconut shell; Phillips et al., 2010). The experimenter then pretended to cut a piece of fruit behind the shell and handed it to the monkey. Monkeys were more surprised and continued searching when the piece handed by the experimenter was a piece of coconut (unexpected), rather than a piece of apple (expected). This finding has been interpreted as showing that monkeys tracked the essential features of objects over events of surface transformations like in the classical verbal (racoon etc.) studies (Keil, 1989). However, it is conceivable that the monkeys interpreted the event shown as mere occlusion (an object of a certain kind disappearing behind another object) or containment relation (an object of a certain kind being placed inside another object) rather than real transformation (where the superficial properties of one and the same object are modified, potentially disguising its kind). Only if monkeys interpreted the event as transformation, the task would entail a test of psychological essentialism, as only then, the monkeys would be potentially ambivalent about the kind of object disappearing into the box (and engage in identity judgements through transformation). It is not fully clear, therefore, whether these findings really provide evidence of essentialist reasoning in non-human primates.

With the present series of experiments, we therefore aimed to study in greater depth the cognitive foundations of psychological essentialism, by testing non-human primates with multiple paradigms, including different types of transformations and a more stringent design. Based on verbal vignettes used in research on psychological essentialism with adults and older children, we presented subjects with events in which objects systematically varied in essential and superficial features. In particular, we adapted one of the classical verbal transformation stories to a non-linguistic format: the surface properties of an object were transformed (e.g., painting the fur of a raccoon), so that the object became superficially more similar to another kind of object (e.g., squirrels), while essentially it remained what it used to be (Keil, 1989). We implemented such transformations in two different established non-verbal tasks, in which subjects have to keep track of and reason about objects: first, a *box task* (e.g., van de Walle et al., 2000), typically used to test object individuation (Experiment 1), and second, an *object choice task* (e.g., Anderson, Sallaberry, & Barbier, 1995), typically used to test object discrimination, object preferences and logical inferences (Experiments 2-4).

2. Experiment 1: The box task

The basic scenario of Experiment 1 was the following: first, apes saw Object 1 (always a piece of banana) being placed into a box; second, they were allowed to search the box, where they found Object 2 (either a piece of banana or a piece of carrot); third, we measured whether they continued searching in the box after retrieving Object 2 (indicating their expectation that there must be another object inside). The apes were presented with five test and two baseline conditions (see Figure 1; see Supplementary material). The five test conditions included differences in kind (essential properties) and/or in surface features between the object placed in the box (Object 1) and the one retrieved from the box (Object 2). In two conditions (as in object individuation studies), Object 1 (the banana piece) was not superficially transformed in any way before being placed in the box, and then apes found either a piece of banana (condition B-B) or a piece of carrot (condition B-C) as Object 2. In the other three conditions (as in the transformation scenarios used in verbal studies on psychological essentialism), Object 1 (the banana piece) underwent a superficial transformation before being placed in the box, being dyed orange to become perceptually more similar to a carrot piece. In the box, subjects found Object 2, which was either the same piece of orange dyed banana (condition DB-DB), a piece of yellow banana (condition DB-B), or a piece of orange carrot (condition DB-C).

Figure 1

In contrast to previous studies, the present set of conditions allowed us to understand whether great apes truly rely on sortal concepts when individuating the number of objects present in the box. If apes engage in true sortal object individuation, they should consider differences in kind (i.e., differences in essential properties) between Object 1 and Object 2 as being more meaningful than mere superficial feature differences. For example, when apes see a piece of banana disappear in the box and then find a piece of carrot (B-C), they should notice the difference in kind, infer that there is still a banana piece in the box and continue searching. However, searching longer in this condition *alone* would not be sufficient to conclude that apes realize that the carrot is not a member of the banana category. Apes might instead rely on superficial features, perceiving “yellowness” going inside the box and “orangeness” coming out of it, and thus searching the box for missing yellowness. Apes’ performance in the condition where the banana is dyed orange before being placed in the box (to become perceptually more similar to a carrot: DB-C) is therefore crucial to understand whether apes perceive the difference between carrot and banana pieces based on essential properties or superficial features. If apes simply rely on superficial features, they should not respond to category change in this condition: they should detect "orangeness" going inside the box and "orangeness" coming out of it, inferring that there was just one object in the box and thus failing to further search the box.

2.1. Methods

Participants. In the 5 test conditions, participants were 14 great apes: 6 bonobos (*Pan paniscus*), 5 orangutans (*Pongo abelii*), and 3 chimpanzees (*Pan troglodytes*), living at the Wolfgang Koehler Primate Research Center at the Leipzig Zoo. The number of subjects depended on the subjects' availability at the research center. We tested the same number of participants for each species also in the 2 baseline conditions, but the identity of few participants differed, because some of them were moved to other facilities. All subjects had prior experience with various cognitive studies. They were tested individually either in their sleeping quarters or in testing rooms, except for mothers with offspring younger than three years, who were tested in presence of their young.

Apparatus and materials. The experimental box was made of opaque plastic (40 x 40 x 34.5 cm) and had a circular opening (approx. 8.5 cm in diameter) on its top middle part, through which the experimenter (E) could insert her hand and forearm to place the stimuli. The box had a false roof (9 cm high from its top part), not visible to the subjects, where a food item could be surreptitiously stored. The frontal part of the box (from the subjects' point of view) had a Plexiglas sliding door which, once opened by E (C.H.), allowed subjects to reach inside the box through a front opening (13 cm wide, 6 cm high). The opening was covered with a curtain made of burlap bag material to prevent subjects from looking inside the box before and while reaching. The experimental box was attached to a Plexiglas panel connected to the testing room. The panel had a circular opening corresponding to the front opening of the box through which the apes could reach inside the box. Slices (1cm thick) of baby bananas (i.e., a small banana variety) and raw carrots served as experimental stimuli. In three conditions (DB-B, DB-B, DB-C) the pieces of banana were dyed with orange liquid food colour. Also carrots were painted orange, to control for the effect of the food colour on the smell/taste of the stimuli.

Design and procedure. All five test conditions, together with a food preference test and a familiarization phase, were carried out on one day (except for two bonobos, who showed a low motivation after the first and fifth experimental trial, and were therefore tested on a second day, after repeating the food preference test). The order of the five test conditions was counterbalanced across individuals. The two baseline conditions were also carried out in one day, after all test conditions were completed. The order of the two baseline conditions was counterbalanced across individuals.

Food preference test: Before the testing conditions, we administered a food preference test to ensure that all apes were able to differentiate the food items and preferred bananas over carrots. The food preference test consisted of four trials, in which subjects had to choose between a piece of carrot and a piece of banana. The side on which food was presented was counterbalanced across individuals and trials. Ninety-three percent of the apes chose the banana slice in all trials.

Familiarization: Before the testing conditions, apes were familiarized with the procedure and the apparatus. E placed a piece of grape on top of the experimental box, made sure that the ape looked and then introduced the grape inside the box. The ape was then allowed to search the box until retrieving the food item. The familiarization phase ended after the ape retrieved the object in three consecutive trials.

Test conditions: Apes were tested in two consecutive sessions that included each of the five test conditions (amounting to a total of two trials per condition and ape). Test conditions are depicted in Figure 1. The object placed in the box (banana) differed in essential properties from the object that was retrieved from it in two conditions: DB-C (i.e., modified banana in - carrot out) and B-C (i.e., banana in - carrot out). In contrast, the same object was placed and also retrieved from the box in two conditions: DB-DB (i.e., modified banana in - modified banana out) and B-B (banana in - banana out). In the fifth condition, DB-B (i.e., modified banana in - unmodified banana out), the object placed into the box and retrieved from it only differed in surface features.

The procedure in the B-B and DB-DB conditions (i.e., when exactly the same object was placed into the box and then retrieved) was as follows. In the B-B condition, E presented a baby banana on top of the box and made sure the ape looked at it. Then E peeled half of it, cut off a slice, showed it to the ape and inserted it into the box. Then E opened the sliding door, so that the ape could search for the piece until it was found (1st search phase). After retrieval, the sliding door was closed and reopened, so that the ape could search for another 20 seconds (2nd search phase). If the ape immediately reached into the box after the retrieval, so that the sliding door could not be closed in the first place, it was closed 20 seconds after the retrieval. The procedure in the DB-DB condition was identical, except that E dyed the banana slice in full view of the ape (with the help of a brush and orange-coloured liquid) before inserting it into the box.

In the DB-C, B-C, and DB-B conditions (i.e., when Object 1 and Object 2 differed in essential properties and/or various degrees of surface properties), the procedure was identical, except that E hid the banana piece on the false roof when inserting it into the box. Before these trials and out of the subject's view, another piece (a carrot or a non-dyed banana) had been hidden on the floor of the box, which the ape could retrieve in the 1st search phase. To prevent any enhancement, however, the same arm and hand movements were done as in the other conditions. In the DB-C and the DB-B conditions, banana slices were first dyed in full view of the ape and then inserted into the box. For each ape, the whole banana was shown and peeled only in the first trial. Thereafter, the same half-peeled banana was used for all other trials in both sessions.

Baseline conditions: After the test conditions, we implemented two baseline conditions (of 2 trials each), to control if apes had a baseline preference, searching longer after finding bananas or carrots. In CtrlB-B condition, the procedure was exactly like in the B-B condition. In the CtrlC-C condition, the procedure was like in the CtrlB-B and B-B conditions, but carrots were used instead of bananas.

Coding and analysis. An assistant filmed the trials focusing on the opening in the box through which the ape reached for the stimuli. The videos were analysed frame by frame using Interact 7 (MANGOLD). The two dependent measures were the duration and the frequency of searching in the second search phase. Searching was defined as introducing the fingers (i.e., at least the second finger joints of the four fingers) into the front opening of the box, while conducting searching movements. We excluded all bouts in which apes only played with the curtain covering the hole, inserted the hand in a supine orientation or simply put the hand into the hole without any further movements. Twenty percent of the video clips were scored by a second observer who was blind to the detailed testing procedure and conditions, as the condition could not be inferred from watching the second search phase only. Moreover the video-camera was placed in such a way that only the hand of the ape and the opening in the box were visible. The reliability between the two observers was very high (for searching duration: Pearson's $r = .998$, $N = 46$, $p < .001$; for searching frequency: Cohen's $k = 1$, $N = 46$, $p < .001$).

For the analyses, we averaged the values of the two trials per condition, for both frequency and duration of searching. In five cases, data for one trial was missing due to a body part of the ape being moved in between the camera and the box opening. In these cases, we took the value of the single measured trial. In one case, data for both trials were missing due to a lack of motivation (i.e, Padana in the DB-DB condition). In this case, no value was entered and the condition was classified as "missing". We conducted two analyses on each of the dependent variables. First, we implemented an ANOVA with "initial food" (modified banana/unmodified banana) and "essential difference" (yes/no) as within-subject factors, to assess whether subjects detected differences in essential properties between the object placed into and retrieved from the box, regardless of the superficial features of the stimuli. Second, an ANOVA with "essential difference" (yes/no) and "surface difference" (yes/no) as within-subjects factors allowed us to directly contrast the impact of differences in essential versus surface properties on the subjects' responses. Finally, we used a t-test to compare searching duration/frequency in the two baseline conditions, to see if there were baseline preferences for one stimulus. All tests were exact and two-tailed, with the α level set at 0.05.

2.2. Results

Given that there were no significant inter-specific differences in performance in any condition ($p \geq .05$ in all cases), we collapsed the data across species. Figure 2 summarizes the mean searching durations and frequencies in all test conditions. An analysis of variance with the two factors initial food (modified banana/unmodified banana) and essential difference (yes/no) revealed a significant effect of essential difference ($F(1,12) = 10.38$, $p = .007$, $\eta_p^2 = .47$), with apes searching longer after finding an object with different essential properties rather than the expected identical object (see Fig. 2a). Neither initial food (*F*(1,12) = .46, p = .508, $\eta_p^2 = .04$) nor the interaction between initial food and essential difference (*F*(1,12) = .008, p = .931, $\eta_p^2 = .00$) showed a significant effect.

Figure 2

A second ANOVA with essential difference (yes/no) and surface difference (yes/no) as factors revealed a significant effect of essential difference (*F*(1,13) = 8.56, p = .012, η_p^2 = .40), with apes searching longer after finding an object with different essential properties rather than an object with different surface features. However, there was no significant effect of surface difference (*F*(1,13) = .18, p = .683, $\eta_p^2 = .01$) or of the interaction between essential and surface difference (*F*(1,13) = .32, p = .583, $\eta_p^2 = .02$).

This pattern was largely confirmed by an analysis of search frequencies (see Fig. 2b). An ANOVA with initial food and essential difference as within-subjects factors revealed a significant effect of essential difference (*F*(1,12) = 13.67, p = .003, $\eta_p^2 = .53$), with apes searching more frequently after finding an object with different essential features rather than the expected identical object. There was no effect of initial food (*F*(1,12) = .53, p = .482, η_p^2 = .04) or of the interaction between initial food and essential difference (*F*(1,12) = .00, p = 1.000, $\eta_p^2 = .00$).

An ANOVA comparing the factors essential difference and surface difference as within-subjects factors revealed a significant effect of essential difference (*F*(1,13) = 4.90, p = .045, $\eta_p^2 = .27$), with apes searching more frequently after finding an object with different essential properties, rather than different surface features. There was no effect of surface difference (*F*(1,13) = .88, p = .365, $\eta_p^2 = .06$), but a significant interaction between essential difference and surface difference (*F*(1,13) = 4.94, *p* = .045, η_p^2 = .28). Post hoc analyses revealed that apes only searched more frequently after finding an object with different surface features if the object essentially remained the same (i.e., within kind; *t*(13) = 2.62, *p* = .021), but not after retrieving an object that differed essentially from the one that was hidden in the box before (i.e., between kind; *t*(13) = .56, *p* = .583).

Finally, there were no differences in searching time (*t*(12) = -.977, *p* = .348) between the two baseline conditions (CtrlB-B: *M* = 2.08, *SE* = .62; CtrlC-C: *M* = 2.56, *SE* = .67). The same was true for searching frequency (*t*(12) = -1.000, *p* = .337), with no differences between the two baseline conditions (CtrlB-B: *M* = .79, *SE* = .24; CtrlC-C: *M* = .96, *SE* = .30). Also, between the CtrlC-C and the C-B conditions there were no differences in searching time (within-fraction of the sample: *t*(8) = -.693, *p* = .508; between-fraction of the sample: *t*(6) = -1.181, *p* = .323) or searching frequencies (within-fraction of the sample: *t*(8) = 1.048, *p* = .325; between-fraction of the sample: *t*(6) = 1.414, *p* = .203).

2.3. Discussion

Great apes individuated and tracked objects as a function of their essential/kind properties, and not just as a function of their surface properties. When observing Object 1 (a piece of banana) disappear in a box and Object 2 exit the box, apes searched the box longer for another missing object when the object retrieved was of a different kind (a piece of carrot), rather than of the same kind (a piece of banana). Crucially, apes did so even when the retrieved object was of a different kind (a piece of carrot), but superficially more similar to the object originally placed in the box (an orange-dyed banana looking more like a carrot than a normal banana). Surface features, in contrast, were largely ignored. If ever, apes used surface features in within-kind comparisons (e.g., DB vs. BB conditions), where they are, in fact, informative.

Therefore, these findings suggest that apes engage in truly sortal object individuation. This corroborates and extends previous studies suggesting that birds (Herrnstein & Loveland, 1964; Herrnstein, Loveland, & Cable, 1976; Pepperberg, 2013) and non-human primates appear capable of a sophisticated form of object individuation that goes beyond mere spatiotemporal or simple featural tracking (Mendes et al., 2008, 2011; Phillips & Santos, 2007; Santos et al., 2002; Uller, Carey, Hauser, & Xu, 1997). In fact, sortal object individuation may be considered a precursor or a primordial form of psychological essentialism, already involving some of the central skills of a more sophisticated essentialist stance, in particular the distinction between deep essential properties that determine identity and merely superficial modifiable surface features. However, the present findings by themselves leave open a more parsimonious alternative: Perhaps there is no categorical distinction between essential and merely superficial features in play. While our findings show that apes did not individuate objects by features in simple ways (relying on total featural identity for individuation), apes might have been operating with more complex feature-based categories (e.g. of a prototype style), in which different features were weighed differently. Since only one type of surface feature was transformed (colour, which was reported to be particularly important to categorize food items, e.g. Shutts, Condry, Santos & Spelke, 2009), apes might have relied on other types of surface features (texture etc.) and weighed these more strongly for individuation.

To rule out such reliance on surface features other than those being transformed, we implemented the transformations in a different task format in the next studies, with a so-called *object choice* test paradigm in which subjects were allowed to choose one out of two food objects. This task was potentially more demanding than the box task in Experiment 1 in several aspects: apes had to simultaneously track the identities of tokens of two categories (e.g., banana and carrot); the transformation events were more extensive, involving modifications along several superficial dimensions of an object (e.g., colour, texture, shape); and these transformations resulted in a radically altered appearance of one of the two objects, so that the two tokens of the two categories (banana/carrot) became perceptually indistinguishable.

3. Experiment 2: The choice task

In Experiment 2 we aimed to test essentialist intuitions in great apes using multidimensional transformations. In this task, we placed a piece of a highly preferred food category (banana) and a piece of a medium preferred food category (carrot) in front of the ape (see Supplementary material). In full view of the ape, we transformed both items to make them superficially look like food items from the same category (e.g., banana). In order to do that, we coated one food item with the peel of the other one and painted its top with the same colour. After the transformation was completed, we let the ape choose the preferred item. We compared the performance in this test condition with a control condition in which apes could not observe the transformation (i.e., the food items were introduced already transformed). Performance in the test condition was also compared to performance in a preference task, in which apes could choose between two untransformed food items (i.e. banana vs. carrot). If apes judged the trans-temporal identity of objects according to essentialism, they should realize that a carrot essentially remains a carrot even if being superficially transformed to look like a banana, sticking to their food preference (e.g., selecting the real banana) even in the face of major transformations.

3.1. Methods

Participants. Participants were 31 great apes: 8 bonobos, 7 orangutans and 16 chimpanzees, living at the Wolfgang Koehler Primate Research Center at the Leipzig Zoo.

Design. All apes were presented on two separate days with a test condition (4 trials) and a control condition (4 trials), in counterbalanced order. Apes also received a food preference condition, once right after the test condition (4 trials) and once right after the control condition (4 trials), to monitor possible changes in food preference. Food preference trials were always run after the test and control condition, not to introduce bananas and carrots beforehand in each session.

Procedure. In the test condition (choice tasks), apes witnessed two types of food manipulations (see Table 1; see Supplementary material): one in which two food stimuli (a raw carrot and a banana slice) were transformed into identical banana-looking stimuli (C-B), and one in which they were transformed into identical carrot-looking stimuli (B-C). Trials of the two conditions were alternated, with 15 participants starting with the C-B and 14 with the B-C condition. In the C-B manipulation, apes faced a table on which E placed two plates, one on the left and the other on the right side. E placed one entire banana behind the first plate and one entire carrot behind the second plate. E cut off a small slice from the banana and gave it to the ape, to make sure that she encoded it in detail. Then E cut off another slice from the banana (approximately 8 mm thick), removed its peel with a knife, placed it on the corresponding plate, and repeated this series of actions with the carrot. Subsequently, E removed everything else from the table (i.e., the previously entire banana and carrot) and started the transformation, by placing one banana peel around each stimulus and painting their top surface yellow. Finally, E pushed the table toward the ape to choose. In the B-C manipulation, the same procedure was followed, but carrot peels were used instead of banana peels, and their top surfaces were painted orange.

In the control condition, E placed two plates on the table, and then mimicked all the movements done in the test condition (i.e., moving the hands and the brush on the empty plates without pieces of food, colour or peels), for approximately the same amount of time required in the test condition. Before pushing the table, E took two pieces of food that had been already manipulated and stored out of the participant's view, placed them sequentially on the table, pushed the table toward the ape and let the ape choose. We counterbalanced the side of the banana/carrot, and the side of the first manipulation. Although we manipulated both stimuli (so that they resulted in two identically looking stimuli), the category-typical appearance was altered only for one of the two stimuli (e.g., the carrot kept the same "carrot-like" colour and texture).

Table 1

Coding and analysis. We considered apes to have made their choice when they pointed to or tried to reach for one of the two food items (very few apes looked intently at one stimulus, instead of pointing). We coded the choice made by each ape, and for each subject and condition we calculated the percentage of trials in which the banana piece was chosen. Twenty percent of the video clips were scored by a second observer who was blind to the detailed testing procedure and conditions (i.e., the second observer had not been previously instructed on the different kinds of conditions of the experiment and on the aims of the study and rated only the final section of the film). The reliability between the two observers for the food item chosen was excellent (Cohen's $k = 1$, $N = 100$, $p < .001$).

We used Wilcoxon tests to compare performance between conditions, and Kruskal-Wallis tests to compare performances across species. When the result of the Kruskal–Wallis test was significant, Mann–Whitney tests were used for pair-wise comparisons (Cohen and Cohen, 1983). We also coded whether apes hesitated (i.e., pointed to both stimuli simultaneously or sequentially, or did not point immediately) before making a choice in the control and test conditions, and we used a Wilcoxon test to compare the percentage of trials in which apes showed hesitation between the two conditions. All tests were exact and two-tailed, with the α level set at 0.05.

3.2. Results

Figure 3 summarizes the mean percentage of trials in which apes chose the banana in the test condition, in the control condition and in the corresponding food preference conditions. Apes selected the banana above chance in the food preference task (Wilcoxon tests; control condition: N = 27, z =-4.835; test condition: N = 29, z =-5.058; p < .001 in both cases). However, they failed to do so in the main task (Wilcoxon tests; control condition: N = 8, z =-.440, p=.800; test condition: N = 9, z =-1.698; p =.119).

Figure 3

Overall, performance did not differ between control and test condition in the main task, with apes selecting the banana in the test condition (i.e., after having witnessed the transformation) as much as in the control condition (Wilcoxon test, N = 13, z =1.452, p =.165). Similarly, their food preference did not differ between control and test conditions (Wilcoxon test, N = 4, z =.604, p =.672). Moreover, their performance reliably differed between food preference tasks and main tasks (Wilcoxon tests; control condition: N = 26, z =-4.468; test condition: N = 22, z =-4.330; p < .001 in both cases). Finally, the percentage of trials in which apes showed hesitation did not differ between the control and the test condition (Wilcoxon test, $N = 6$, $z =-.045$, $p = .964$).

Performance reliably differed across species in the test condition (Kruskal-Wallis test, $\chi^2 = 10.093$, df = 2, $N = 31$, $p =.003$; Fig. 4). Pairwise tests revealed that bonobos selected the banana more often than chimpanzees (Mann-Whitney-U test, $N = 16$, $z = -3.188$, $p =.001$) and orangutans (Mann-Whitney-U test, $N = 8$, $z = -2.170$, $p = .034$). Furthermore, bonobos, unlike chimpanzees and orangutans, selected the banana significantly above chance (Wilcoxon test, $N = 6$, $z =-2.251$, $p = .031$). There were also inter-specific differences in the preference task associated to the test condition (Kruskal-Wallis test, $\chi^2 = 8.925$, df = 2, $N = 31$, $p = .009$; Fig. 4). Pairwise tests revealed that bonobos selected bananas less often than chimpanzees (Mann-Whitney-U test, $N = 16$, $z = -2.748$, $p = .007$). In contrast, there were no inter-specific differences in the control condition (Kruskal-Wallis test, $\chi^2 = .147$, df = 2, $N = 30$, $p =.936$; Fig. 4) and in the preference task associated to this condition (Kruskal-Wallis test, $\chi^2 = 3.328$, df = 2, $N = 30$, $p = .221$, Fig. 4). Finally, there was no correlation between performance in the test condition and in the corresponding food preference condition (Pearson's $r = -.276$, $N = 31$, $p = .133$).

Figure 4

3.3. Discussion

Apes reliably preferred the banana in the food preference tasks of the control and test conditions, but in general failed to reliably track and select the banana piece after manipulations were implemented (i.e., changing the appearance of the banana into a carrot, or vice versa). In fact, after having witnessed the transformations in the test condition, they performed just like in the control condition, where they were confronted with the two perceptually indistinguishable objects without having seen their transformations. With the exception of the bonobos, who seemed to track the banana piece through transformations and reliably selected it in the test condition (see below), there was thus no evidence for apes' distinction between real and apparent kinds in accordance with psychological essentialism.

These overall null findings in the object choice tasks contrast sharply with the positive findings in the box task in Experiment 1. Why might this be the case? One possibility is that the more stringent task in Experiment 2 should be considered the valid test and the findings in Experiment 1 might thus be false positives. Alternatively, the findings with the current version of the object choice task in Experiment 2 might be false negatives, masking apes' competence due to extraneous performance factors. In particular, the object choice task used here introduced an even higher level of representational and executive task demands compared to the box task.

First, judging identity through transformations necessarily includes certain representational demands. Specifically, apes had to build and compare multiple models (past-present) of the same entity in order to track the objects through the transformation and judge their category membership. Doing so required apes to engage in advanced forms of reasoning based on comparing past and present models of an entity whose transformations created an appearance-reality conflict (e.g., between what kind of object the transformed entity really was, and what kind it looked like after the transformation). Solving appearance-reality conflicts is cognitively demanding as only a few chimpanzees seem able to do it (Krachun, Call & Tomasello, 2009). However, in contrast to the box task (using very basic transformations, i.e. altering just one surface feature), the choice task exacerbated the representational challenge (i) by including very extensive object transformations (i.e., across more dimensions, such as colour, texture and shape), resulting in fully changed appearances, and (ii) by changing the outer appearance of two objects simultaneously, only one of which maintained the properties pertaining to its real category membership after the transformation.

Second, the choice task was more challenging than the box task also in terms of general executive demands (i.e., overall memory and inhibitory demands). In the choice task, for example, apes had to simultaneously process and compare information referring to two different objects: they had to monitor *two* trajectories and remember past (real) and present (apparent) category membership of *two* objects. Therefore, the choice task was clearly more demanding in terms of working memory and processing demands (e.g., De Loache, 2000; Deak, 2006). Limits in working memory can result in low performance in a variety of tasks, despite individuals possessing all the cognitive skills otherwise needed to solve the task. In a study testing apes' ability to use tools, for instance, failures to select the right tool mainly depended on individuals' limits in memory (or attention), rather than on a failure to represent connection in tools (Seed, Seddon, Greene & Call, 2012). Further, in the choice task apes had to choose between two identical objects presented in full view. The presence of visible features at the time of choice may have overridden any other information that they could have used to disambiguate the food items. Both stimuli were desirable, and selecting one of the two might have been harder. Thus, the choice task might have also imposed higher inhibitory demands than the box task. Interestingly, the better performance of bonobos in this Experiment also fits with this hypothesis. In particular, bonobos were the only species reliably tracking food through transformations, but also the one showing a less marked preference for bananas over carrots, as compared to chimpanzees and orangutans (for similar results with a different set-up, see Sanchez, Pereto & Call, in press). Such a reduced preference might have allowed them to inhibit false selections more easily, and consequently perform better overall – a finding that would be consistent with the Yerkes-Dodson law, according to which performance might be better with an intermediary state of physiological arousal (Yerkes and Dodson, 1908). Low preferences (leading to low arousal and failure to be motivated) and very high preferences (leading to high arousal and inhibitory failures) may thus both result in lower performance, as compared to medium preferences (establishing the ideal blend of motivation and cognitive control).

Third, another potentially complicating factor of the object choice task is motivation, relating to the task's payoff structure: the more extensive transformations in the choice task clearly resulted in a much less attractive object-to-cover ratio than in the box task (i.e., a relatively large proportion of undesired cover materials in relation to the desired banana content). Apes might have simply been less motivated to select the preferred stimulus after the extensive transformations (resulting in two moderately interesting items). Given that executive demands and motivational aspects may have masked subjects' true representational capacities, we addressed the impact of each of these factors on performance in the next experiment. More specifically, we investigated whether apes would be able to deal with the representational demands of the task provided working memory and inhibitory control demands were reduced (see Experiment 3, inhibition and memory controls) and motivation was increased (see Experiment 3, ratio control). We reduced working memory load by eliminating one of the food pieces and presenting only one type of transformation (i.e., banana transformation). We reduced inhibitory demands by covering the available alternatives. We increased motivation by increasing the size of the food item core (hidden food item) in relation to its surface (cover materials).

4. Experiment 3: Choice controls

In order to clarify whether apes' poor performance in Experiment 2 reflects a genuine limitation in essentialist reasoning, or might have been due to performance factors, Experiment 3 implemented a number of follow-up and control conditions of the general setup of the object choice task used in Experiment 2, in which memory and inhibition were reduced and motivation increased. If poor inhibitory control and/or reduced working memory capabilities were responsible for the negative results in Experiment 2, reducing memory and inhibitory demands would improve performance (inhibition control, memory control). Similarly, if an insufficient incentive was responsible for the results of Experiment 2, increasing the incentive (by upgrading the ratio between banana and painting/peel) would also improve performance (ratio control).

4.1. Methods

Participants. Participants were 29 great apes: 8 bonobos, 6 orangutans and 15 chimpanzees, living at the Wolfgang Koehler Primate Research Center at the Leipzig Zoo. The participants were exactly the same ones tested in Experiment 2, with the exception of one orangutan (Batak), who could not be tested alone, and one chimpanzee (Ulla) was unavailable at the time testing took place.

Procedure. We administered one inhibition condition (with the corresponding control), followed by one memory condition (with the corresponding control), and then by one ratio condition (with the corresponding control), each followed by a food preference task (see Table 1).

Inhibition task. The procedure was exactly like in Experiment 2, but stimuli were occluded right after being transformed and were never shown simultaneously in the test and control trials (see Table 1; see Supplementary material). Given that both stimuli were desirable, reducing the food saliency might have imposed lower inhibitory demands and might have thus made the selection easier for the apes.

Memory task. The procedure was identical to Experiment 2, with the exception that a balled-up paper (instead of a carrot) was transformed into a banana, by adding a banana peel and painting it (see Supplementary material). Therefore, apes only witnessed one type of manipulation involving only one food kind, reducing the memory load.

Ratio task. The procedure was identical to Experiment 2, with the exception that slices of banana were 3 cm instead of 0.8 cm thick (see Supplementary material). In contrast to the other conditions, transformed stimuli were placed into little transparent tubes made of plastic wrap, in order to stabilize them and avoid them falling. By using larger stimuli (i.e., visibly increasing the ratio between the content and the peel/painting), subjects might have been more motivated to select the banana. Moreover, wrong selections became more costly, because mistakes cost apes larger banana pieces.

Twenty percent of the video clips were scored by a second observer who was blind to the detailed testing procedure and conditions (i.e., the second observer had not been previously instructed on the conditions of the experiment and on their aim and rated only the final section of the film). The reliability between the two observers for the food item chosen was excellent (Cohen’s $k = 1$, $N = 277$, $p < .001$).

4.2. Results

Figure 5 summarizes the mean percentage of trials in which apes chose the banana in the test and control trials of the inhibition, memory and ratio conditions, as well as performance in the corresponding food preference trials, and in Experiment 2 (for comparison). Performance did not reliably differ across species (Kruskal-Wallis test, *p* > .05 in all cases). Apes selected the banana more than chance across food preference tasks, in all test and control conditions (Wilcoxon test, *p* < .001 in all cases). Further, they reliably selected the real banana in the memory and the ratio test conditions (Wilcoxon test, memory: *N* = 17, *z* =-3.400; ratio: *N* = 23, *z* =-4.119; *p* < .01 in both cases), but failed to do so in the inhibition test condition (Wilcoxon test, *N* = 14, *z* = -1.784, *p* > .05). As expected, apes also failed to select the real banana in all control conditions (Wilcoxon test, *p* > .05 in all cases).

Figure 5

The percentage of trials in which apes chose the banana reliably differed between control and test conditions in the memory and the ratio tasks (Wilcoxon test, memory*:* *N* = 18, *z* =-3.203, ratio: *N* = 20, *z* =-3.871, *p* ≤ .001 in both cases), but not in the inhibition task (Wilcoxon test, *N* = 9, *z* = -.996, *p* = .348). Further, the percentage of trials in which apes chose the banana in the food preference tasks never differed between control and test conditions (Wilcoxon test, *p* > .05 in all cases). Finally, there was no correlation between performance in the test condition and the corresponding food preference condition in the memory, ratio and inhibition tasks (*p* > .05 in all cases).

4.3. Discussion

Apes reliably selected the banana in the memory and ratio test conditions, but not in the inhibition one. Importantly, apes failed to select the preferred banana in all controls (where they did not monitor the transformation), confirming that it was not possible to identify the real banana on the basis of perceivable cues. In contrast to Experiment 2, we also found no inter-specific differences in performance. Overall these findings suggest that apes generally appreciate that objects keep their identity through transformations, but representational and/or executive task demands might have masked this competence in Experiment 2. More specifically, the fact that performance increased in the memory task (but not in the inhibition task) further suggests that the high load on working memory (rather than inhibitory problems) might have been responsible for the apes' failure in Experiment 2.

The findings of Experiment 3 thus seem to suggest that apes can successfully track objects as a function of their essential kind properties, and that failure to do so mainly depends on memory and motivational issues. Whereas reducing inhibitory demands had no effect on performance, reducing the memory load and increasing the food incentive both increased performance. However, there is one potential problem with this interpretation: given that the tasks in the present experiment were administered in fixed order (inhibition-memory-ratio), the fact that performance was better in the memory and ratio conditions could merely reflect a learning effect. We therefore designed Experiment 4, in order to rule out this alternative explanation (i.e., apes simply learned how to solve the task over time).

5. Experiment 4: Learning effects?

In order to rule out that success in Experiment 3 depended on apes having learned the contingencies of the tasks, we repeated the choice task (which was the first condition administered in Experiment 2), and the ratio task (which was the last condition administered in Experiment 3), following exactly the procedures in Experiment 2 and 3, with the same sample of subjects, and with counterbalanced order of conditions. The logic was the following: if apes had learned to solve the object choice task during the course of Experiment 3, they should have performed at equally proficient levels in the choice and ratio tasks in Experiment 4. In contrast, if the two tasks differ substantially, so that the ratio task is *per se* easier for apes, given its reduced motivational task demands, apes should perform in the ratio task like they did in Experiment 3, and in the choice task like they did in Experiment 2.

5.1. Methods

Participants. Participants were 27 great apes: 6 bonobos, 6 orangutans and 15 chimpanzees, living at the Wolfgang Koehler Primate Research Center at the Leipzig Zoo. The participants were exactly the same ones tested in Experiment 3, with the exception of two bonobos (Jasongo and Joey), who could not be tested because of conflicts in the group, one chimpanzee (Brigitte), who died before this task, and was replaced by another (Alexandra).

Procedure. We alternated the choice task (as done in Experiment 2: choice2) and the ratio task (as done in Experiment 3: ratio2) in a counterbalanced order. For each individual, we administered only one condition per day, following exactly the same procedure as in Experiments 2 and 3. Twenty percent of the video clips were scored by a second observer who was blind to the detailed testing procedure and conditions (i.e., the second observer had not been previously instructed on the conditions of the experiment and on their aim and rated only the final section of the film). The reliability between the two observers for the food item chosen was excellent (Cohen’s $k = 1$, $N = 176$, $p < .001$).

5.2. Results

Figure 6 summarizes the mean percentage of trials in which apes chose the banana in the test and control trials of the choice2 and ratio2 conditions, as well as performance in the corresponding food preference trials, and in Experiment 2 (for comparison). Performance did not reliably differ across species, in any task and condition (Kruskal-Wallis test, $p > .05$ in all cases). Apes selected the banana more often than expected by chance across all food preference tasks (Wilcoxon test, $p < .001$ in all cases). Moreover, they reliably selected the real banana in the ratio2 test condition (Wilcoxon test, $N = 17$, $z = -3.400$, $p < .01$), but not in the choice2 test condition (Wilcoxon test, $N = 12$, $z = -1.576$; $p = .142$). As expected, they selected at chance level in the control conditions of both tasks (Wilcoxon test, $p > .05$ in both cases).

Figure 6

The percentage of trials in which apes chose the banana reliably differed between control and test conditions in the ratio2 task (Wilcoxon test, $N = 16$, $z = -3.098$, $p = .001$), but not in the choice2 task (Wilcoxon test, $N = 11$, $z = -.734$, $p >= .496$). Further, the percentage of trials in which apes chose the banana in the food preference tasks never differed between control and test conditions (Wilcoxon test, $p > .05$ in both cases).

5.3. Discussion

Once again, great apes reliably selected the banana in the ratio2 but not in the choice2 test condition. If great apes in Experiment 3 had simply learned how to solve the tasks, they would have been successful also in the choice2 test condition, which was not the case. Instead, great apes reliably tracked bananas in the task offering higher incentives (ratio2 test condition), but not in the task lacking them (choice2 test condition). These results confirm that apes can successfully track objects through transformations, provided (i) that they are strongly motivated to do that, despite high executive demands (as in the ratio task of Experiments 3 and 4), or (ii) that the memory load is sufficiently reduced, even if motivation is lower (as in the memory task of Experiment 3).

6. General Discussion

The present experiments were designed to investigate the cognitive foundations of great apes' representation and individuation of objects. To this end, we devised non-verbal tasks that were modelled on classical verbal transformation scenarios, which are widely used in research on psychological essentialism with human children and adults. In Experiment 1, we used a modified box task commonly used in object individuation research. Subjects saw an object (Object 1) disappear in the box, but retrieved a different object (Object 2) from it, and we measured whether subjects continued searching inside the box. We systematically varied whether Object 1 and 2 differed in kind and/or in superficial properties. Subjects based their object individuation (indicated by the amount of time they spent searching after retrieving Object 2) on kind differences between Object 1 and 2, largely ignoring superficial differences. In Experiments 2-4, a different type of task was used, in which subjects first saw two tokens of different food kinds (e.g. banana vs. carrot slice), one of which was then radically transformed (concerning colour, shape etc.) so that the two became perceptually indistinguishable. When given a choice between the two objects, apes in Experiment 2 failed to discriminate them. The subsequent Experiments 3 and 4, however, suggested that this negative finding reflects a performance rather than a competence deficit: once the task was modified to reduce memory load and increase the incentive, subjects performed successfully, tracking and choosing the more attractive object (banana slice) over superficial transformations, even though the preferred object became perceptually indistinguishable from the other object.

These findings add to previous research suggesting that great apes' object individuation is not confined to spatiotemporal tracking. If apes had been merely sensitive to spatiotemporal information, they would have been unable to solve the task in Experiment 1, and the task in Experiment 2 should have been utterly trivial – yet, the pattern of results was the reverse. This provides evidence that apes were not considering manipulations as simple occlusions, but rather as real transformations. More importantly, the present findings go beyond existing research, by showing that apes' object individuation is not confined to tracking superficial feature information either. Previous research only offered indirect and inconclusive evidence for the claim that nonhuman primates use sortal concepts to individuate objects (see Xu, 2007). In all these studies, the participants individuated normal objects whose kind (essential properties) and appearance (surface properties) were necessarily confounded, so that it remained unclear whether apes' object individuation was based on tracking surface features or kind (essential features). By using complex transformation scenarios that disentangled essential and mere surface differences, our study provides compelling evidence that great apes do indeed engage in true sortal object individuation. In contrast to influential philosophical (Quine, 1960) and psychological claims (Xu, 2002), sortal object individuation thus clearly antedates the evolution of language.

In this study, apes distinguished between deeper properties of an object and merely superficial features, relying on the former while neglecting the latter when judging category membership and identity over time. They thus performed much like adults and older children in verbal transformation scenarios, where subjects base their essentialist judgments of identity on kind information (a racoon is a racoon is a racoon…), while largely disregarding superficial feature changes (painting the racoon like a squirrel etc.) (Keil, 1989). In rudimentary form, the sortal object individuation documented here in great apes can be seen as a kind of psychological essentialism, perhaps the phylogenetically and ontogenetically most basic one (Rakoczy & Cacchione, 2014).

Characterizing this primordial form of psychological essentialism more precisely will be a central challenge for future research in comparative and developmental cognitive science. First of all, while showing that apes are capable of this form of cognition, Experiment 2 also showed that the use of this capacity can be fragile. Future research should therefore explore more systematically the scope and limitations of such a basic form of essentialist reasoning. Second, how widespread might such capacities be in the primate lineage or beyond? One of the most exciting questions for the future concerns potential convergent evolution of such cognitive abilities even beyond the primate lineage. Recent work has shown some basic capacities for object individuation according to properties in dogs (Bräuer & Call, 2011), and even in newly born chicks according to properties, spatiotemporal and kind information (Fontanari, Rugani, Regolin, & Vallortigara, 2011, 2014). It remains to be clarified whether sortal object individuation and basic forms of psychological essentialism are even further widespread in the animal kingdom. Third, how does sortal object individuation, as documented here, relate to the more complex forms of psychological essentialism that have been extensively studied in older children and adults? In human infants, sortal object individuation has been amply documented from around one year of age (see e.g., Needham & Baillargeon, 2000; Krøjgaard, 2004; Xu, 2005). From around age 4, children develop more comprehensive and complex forms of psychological essentialism that are thought to underlie a suite of cognitive processes, such as categorization, inductive generalisation, similarity ratings, appearance-reality distinctions, causal reasoning, and intuitive theory building (Gelman & Wellman, 1991; Gottfried & Gelman, 2005; Medin, 1989). So far it remains unclear, from a developmental point of view, how children get from the more primitive to the more complex forms of essentialism, and from a comparative point of view, how much apes (and other species) might share of the more complex types of psychological essentialism that we find in human adults.

References

Anderson, J. R., Sallaberry, P. & Barbier, H. 1995. Use of experimenter-given cues during object-choice tasks by capuchin monkeys. *Animal Behaviour, 49*, 201-208.

Bräuer, J., & Call, J. (2011). The magic cup: great apes and domestic dogs individuate objects according to their properties. *Journal of Comparative Psychology*, *125*, 353-361.

Cacchione, T., Schaub, S., & Rakoczy, H. (2013). Fourteen-month-old infants infer the continuous identity of objects on the basis of non-visible causal properties. *Developmental Psychology. 49*, 1325-1329.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum Associates Inc.

DeLoache, J. (2000). Dual representation and young children's use of scale models. *Child Development*, *71*, 329-338.

Deak, G. O. (2006). Do children really confuse appearance and reality? *Trends in Cognitive Sciences*, *10*, 546-550.

Fontanari, L., Rugani, R., Regolin, L., & Vallortigara, G. (2011). Object individuation in 3-day-old chicks: use of property and spatiotemporal information. *Developmental Science*, *14*, 1235-1244.

Fontanari, L., Rugani, R., Regolin, L., & Vallortigara, G. (2014). Use of kind information for object individuation in young domestic chicks. *Animal Cognition, 17,* 925-935.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23,* 183-209.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: the role of categories and appearances. *Child Development, 58,* 1532-1541.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: early understandings of the non-obvious. *Cognition, 38,* 213-244.

Gottfried, G. M., & Gelman, S. A. (2005). Developing domain-specific causal-explanatory frameworks: the role of insides and immanence. *Cognitive Development, 20,* 137-158.

Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science, 146,* 549-551.

Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 2,* 285-302.

Keil, F. C. (1982). Intelligence and the rest of cognition. *Intelligence, 6,* 1-21.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* Cambridge, MA: The MIT Press.

Krachun, C., Call, J., & Tomasello, M. (2009). Can chimpanzees distinguish appearance from reality? *Cognition, 112,* 435-450.

Kripke, S. (1972). Naming and necessity. In D. Davidson & G. Harman (Eds.), *Semantic of natural language*. Dordrecht, Holland: Reidel.

Krøjgaard, P. (2004). A review of object individuation in infancy. *British Journal of Developmental Psychology, 22,* 159-183.

Medin, D. (1989). Concepts and conceptual structure. *American Psychologist, 44,*1469-1481.

Medin, D., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-195). Cambridge, MA: Cambridge University Press.

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

Mendes, N., Rakoczy, H., & Call, J. (2011). Primates do not spontaneously use shape properties for object individuation: a competence or a performance problem? *Animal Cognition, 14,* 407-414.

Needham, A., & Baillargeon, R. (2000). Infants' use of featural and experiential information in segregating and individuating objects: a reply to Xu, Carey, and Welch (1999). *Cognition, 74,* 255-284.

Newman, G. E., Herrmann, P., Wynn, K., & Keil, F. C. (2008). Biases towards internal features in infants' reasoning about objects. *Cognition, 107,* 420-432.

Pepperberg, I. M. (2013). Abstract concepts: data from a Grey parrot. *Behavioural Processes, 93,* 82-90.

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.

Phillips, W., Shankar, M., & Santos, L. R. (2010). Essentialism in the absence of language? Evidence from rhesus monkeys (*Macaca mulatta*). *Developmental Science, 13,* F1-F7.

Putnam, H. (1975). The meaning of meaning. In H. Putnam (Ed.), *Mind, language and reality*. London: Cambridge University Press.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Rakoczy, H., & Cacchione, T. (2014). The developmental and evolutionary origins of psychological essentialism lie in sortal object individuation. *Behavioral and Brain Sciences, 37,* 500-501.

Sanchez, A., Pereto, M. & Call, J. (in press). Differences in between-reinforcer value modulate the selective value effect in great apes (*Pan troglodyes, P. paniscus, Gorilla gorilla, Pongo abelii*). *Journal of Comparative Psychology.*

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.

Seed, A., Seddon, E., Greene, B., & Call, J. (2012), Chimpanzees ‘folk physics’: bringing failures into focus. *Philosophical Transactions of the Royal Society B, 367,* 2743-2752.

Shutts, K., Condry, K. F., Santos, L. R., & Spelke, E.S. (2009). Core knowledge and its limits: The domain of food. *Cognition, 112,* 120-140.

Uller C., Carey S., Hauser M., & Xu F. (1997). Is language needed for constructing sortal concepts? A study with nonhuman primates. In *Proceedings of the 21st Annual Boston University Conference on Language Development* (pp. 665-677). Somerville, MA*:* Cascadilla Press.

van de Walle, G. A., Carey, S., & Prevor, M. (2000). Bases for object individuation in infancy: evidence from manual search. *Journal of Cognition and Development, 1,* 249-280.

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

Xu, F (2005). Categories, kinds, and object individuation in infancy. In L. Gershkoff-Stowe, D. Rakison (Eds.) *Building object categories in developmental time* (pp. 63-89). Lawrence Erlbaum.

Xu, F. (2007). Language acquisition and concept formation: count nouns and object kinds. In G. Gaskell (Ed.), *Oxford Handbook of Psycholinguistics* (pp. 627-634). Oxford University Press.

Xu, F., & Baker, A. (2005). Object individuation in 10-month-old infants using a simplified manual search method. *Journal of Cognition and Development, 6*, 307-323.

Xu, F., Carey, S., & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive Psychology, 49*, 155-190.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18,* 459-482.

Figure and table captions

Table 1. *Design of Experiments 2-4.*

Figure 1. Test stimuli entering and exiting the box in the five conditions tested (B-C, DB-C, DB-B,DB-DB, B-B). B indicates “banana”; C indicates “carrot”; DB indicates “orange dyed banana”; paint-brush symbol indicates that the banana was dyed orange before being placed into the box, in full view of the ape.

Figure 2. Mean (a) duration of searching (+SE) and (b) frequency of searching (+SE) in the test conditions (B-C, DB-C, DB-B, DB-DB, B-B).

Figure 3. Mean percentage (+SE) of banana-choices in the Control and Test conditions, and in the corresponding Food preference conditions. Only in the Test condition could apes monitor the manipulation of the stimuli. The black line represents chance performance.

Figure 4. Mean percentage (+SE) of banana-choices in the Control and Test conditions, and in the corresponding Food preference conditions, plotted for all and for species separate. Only in the Test condition could apes monitor the manipulation of the stimuli. The black line represents chance performance.

Figure 5. Mean percentage (+SE) of banana choices in the control and test trials of the inhibition, memory and ratio conditions, and in the corresponding food preference trials. The results from Experiment 2 are added for comparison. Only in the test trials could apes monitor the manipulation of the stimuli. The black line represents chance performance.

Figure 6. Mean percentage (+SE) of banana choices in the control and test trials of the choice2 and ratio2 conditions, and in the corresponding food preference trials. The results from Experiment 2 are added for comparison. Only in the test trials could apes monitor the manipulation of the stimuli. The black line represents chance performance.

Figure 1

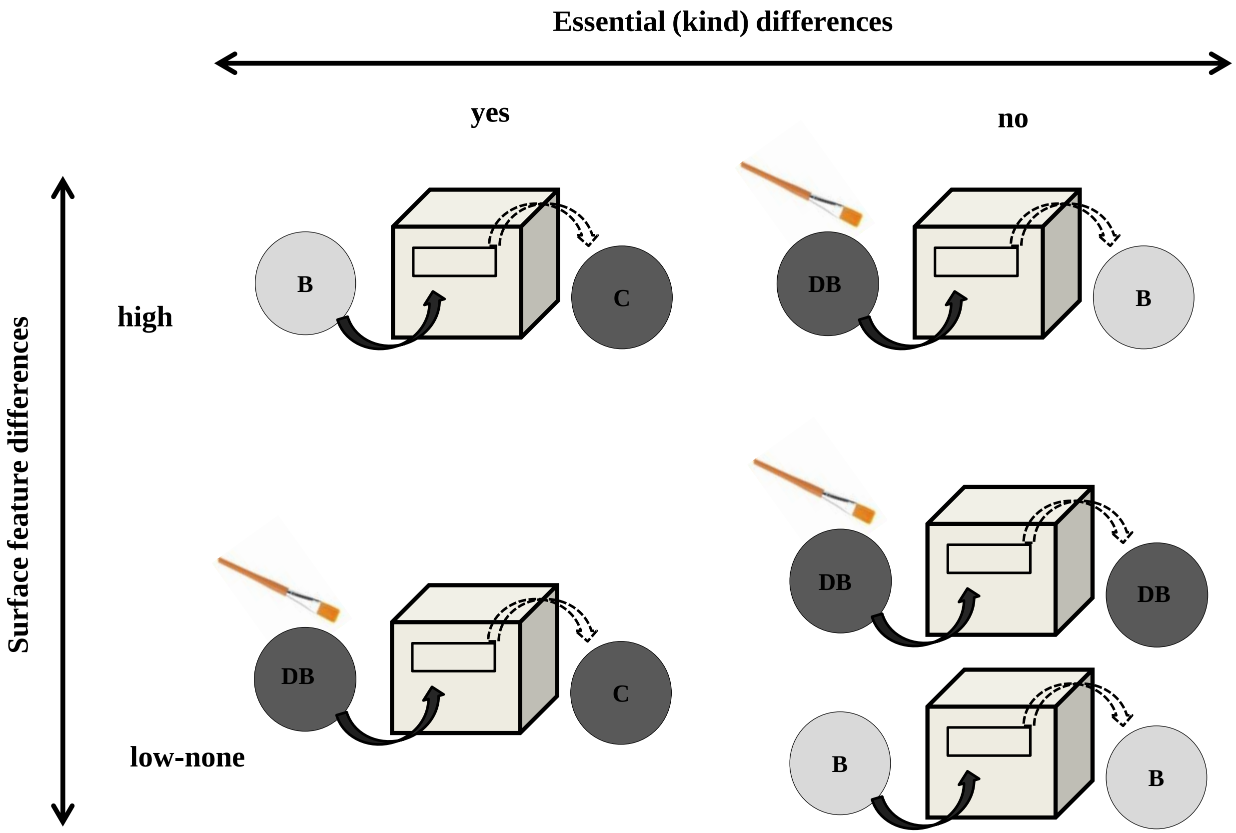

Essential (kind) differences
yes
no
Surface feature differences
high
low-none
B
C
DB
B
DB
C
DB
DB
B
B

Figure 2

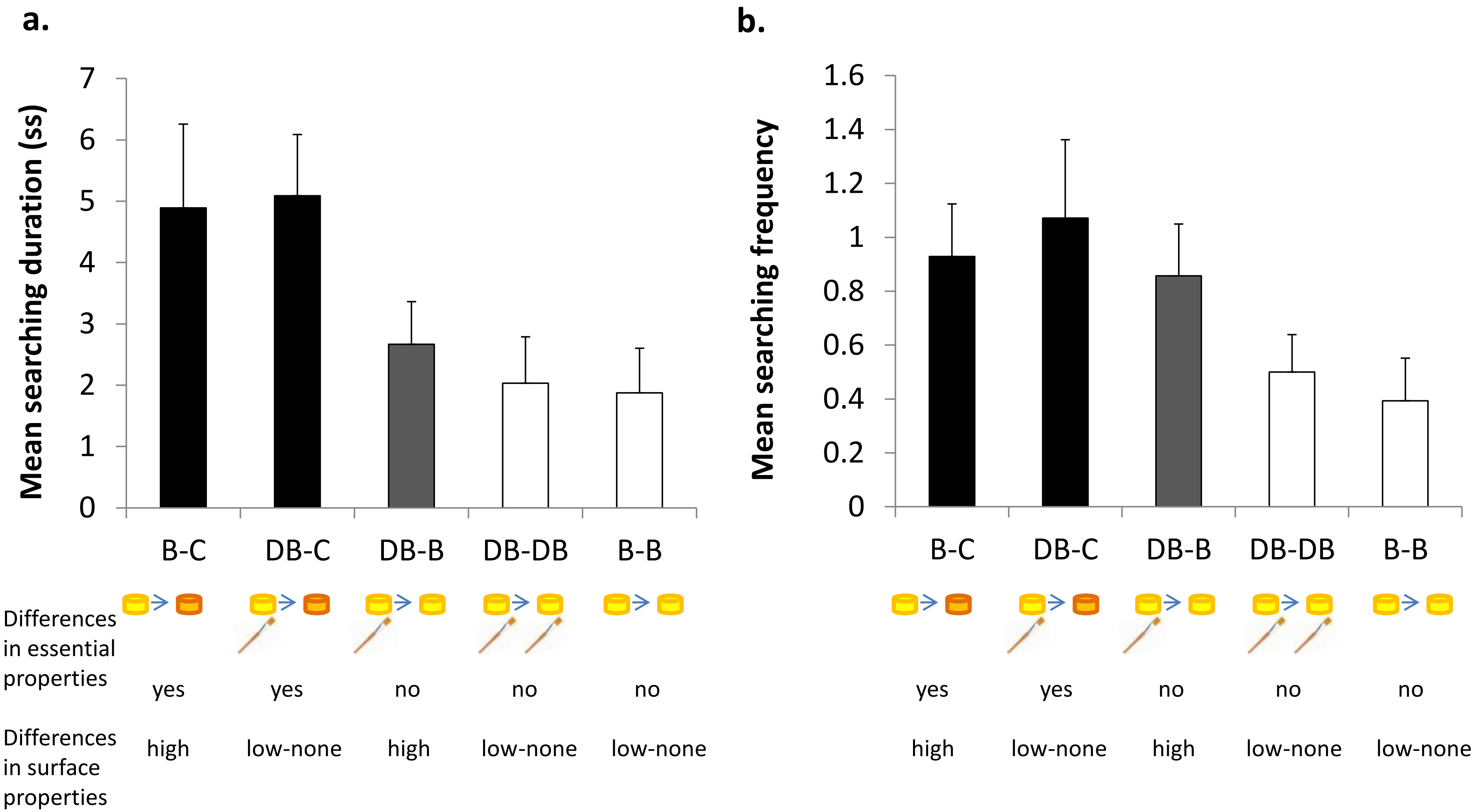

a.
Mean searching duration (ss)
0
1
2
3
4
5
6
7
B-C
DB-C
DB-B
DB-DB
B-B
b.
Mean searching frequency
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
B-C
DB-C
DB-B
DB-DB
B-B
Differences in essential properties
yes
yes
no
no
no
yes
yes
no
no
no
Differences in surface properties
high
low-none
high
low-none
low-none
high
low-none
high
low-none
low-none

Figure 3

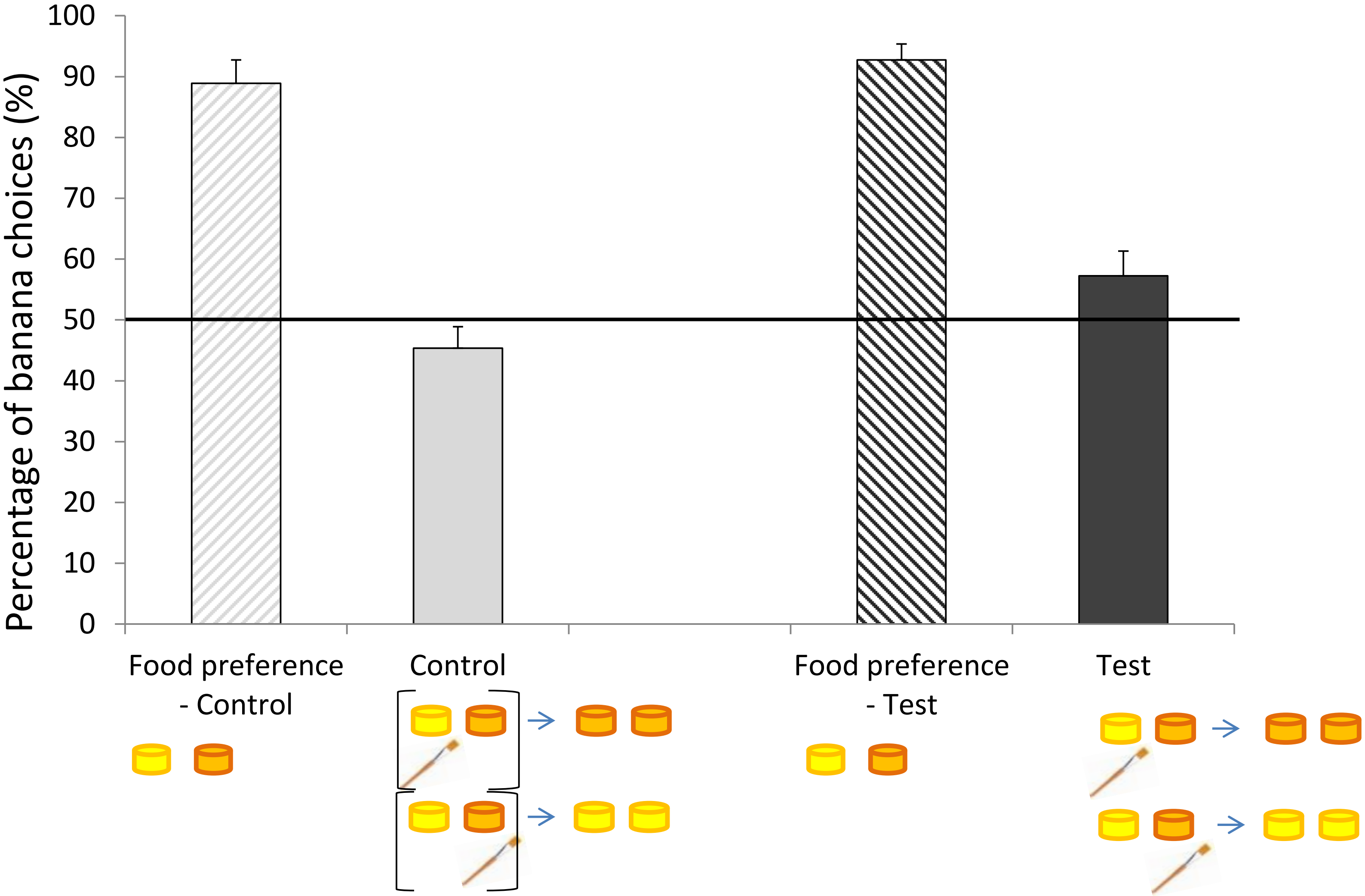
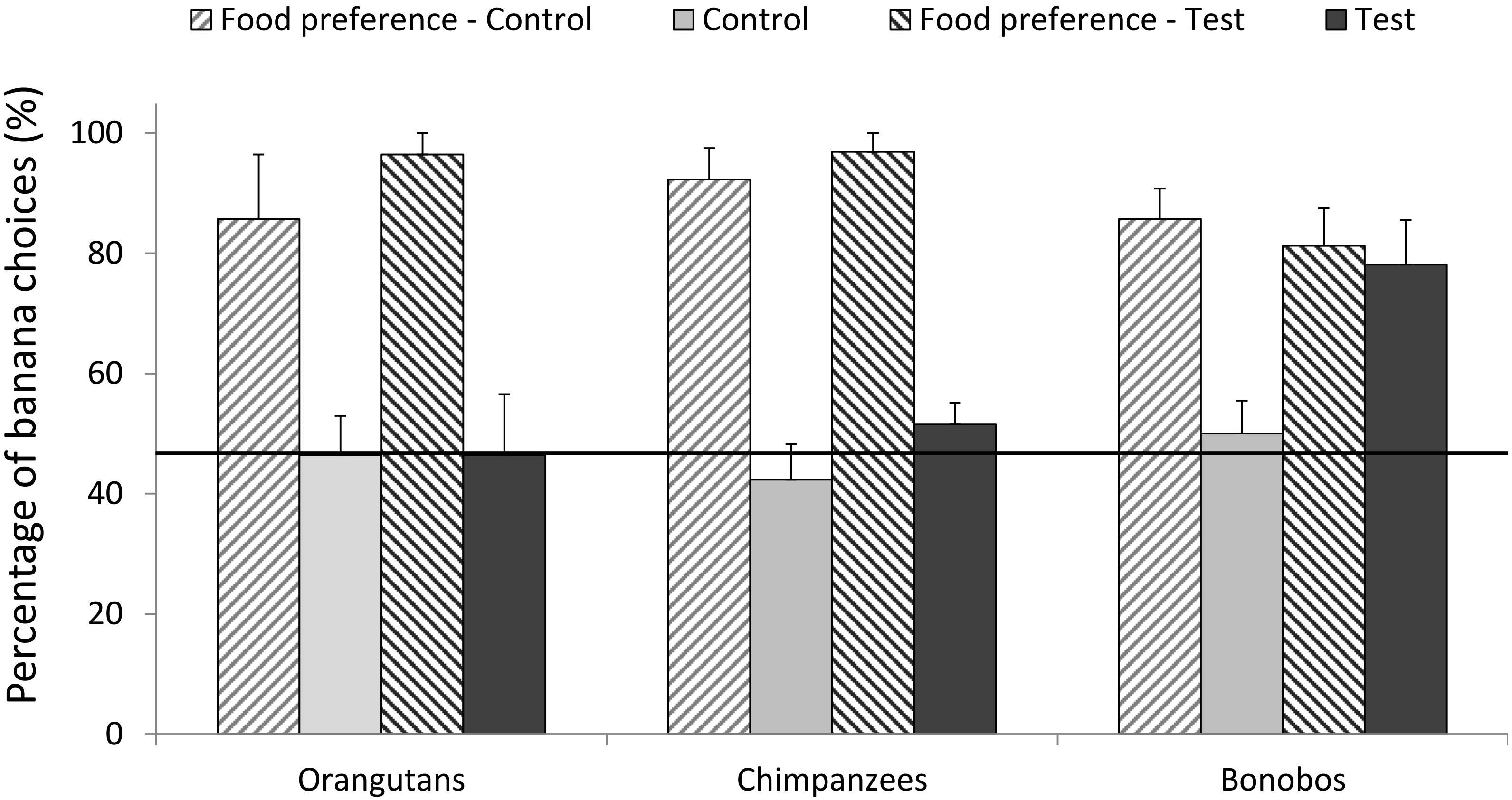

Percentage of banana choices (%)
100
90
80
70
60
50
40
30
20
10
0
Food preference - Control
Control
Food preference - Test
Test

Figure 4



Food preference - Control
Control
Food preference - Test
Test
Percentage of banana choices (%)
0
20
40
60
80
100
Orangutans
Chimpanzees
Bonobos

Figure 5

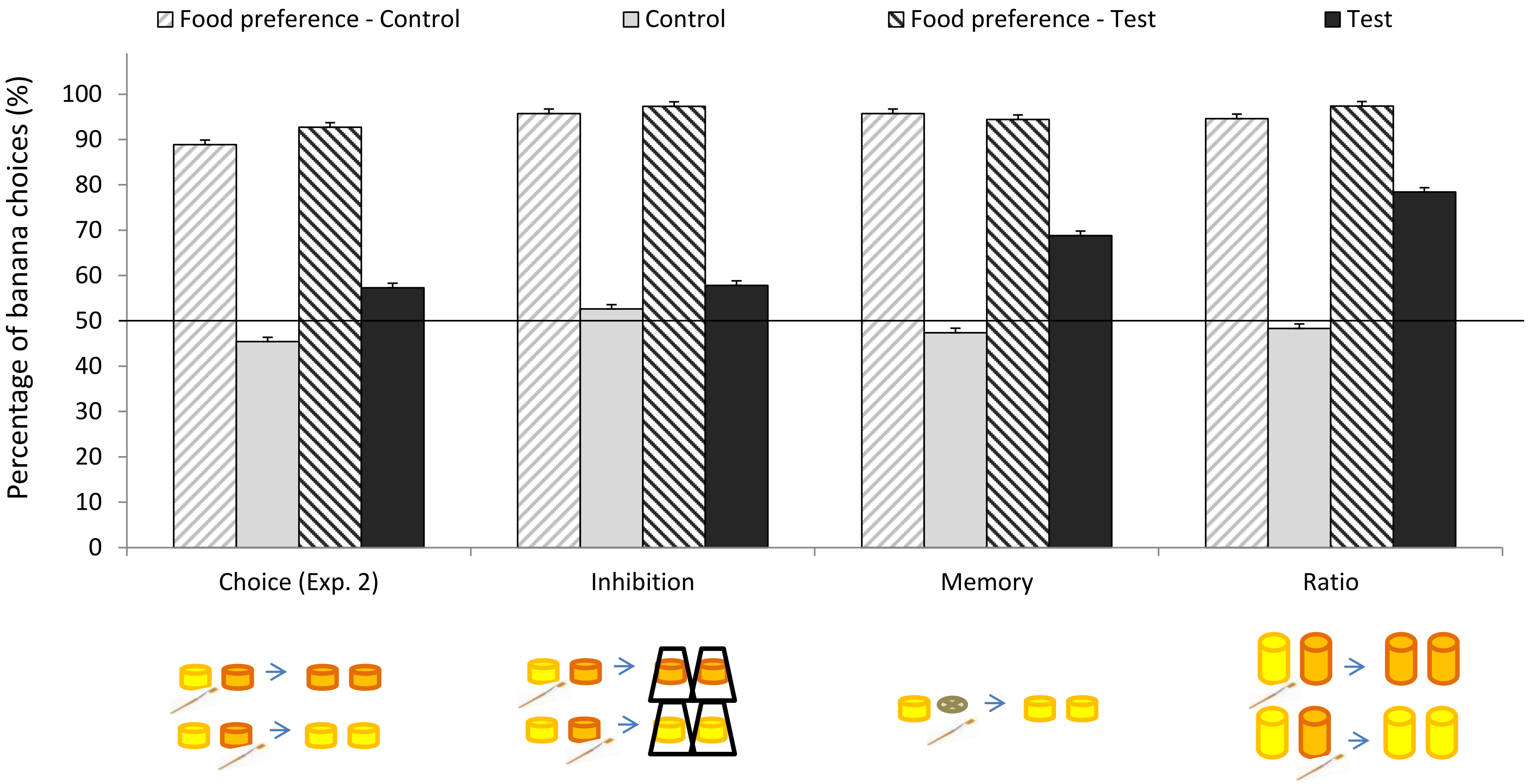

▨ Food preference - Control
☐ Control
▧ Food preference - Test
■ Test
Percentage of banana choices (%)
100
90
80
70
60
50
40
30
20
10
0
Choice (Exp. 2)
Inhibition
Memory
Ratio

Figure 6

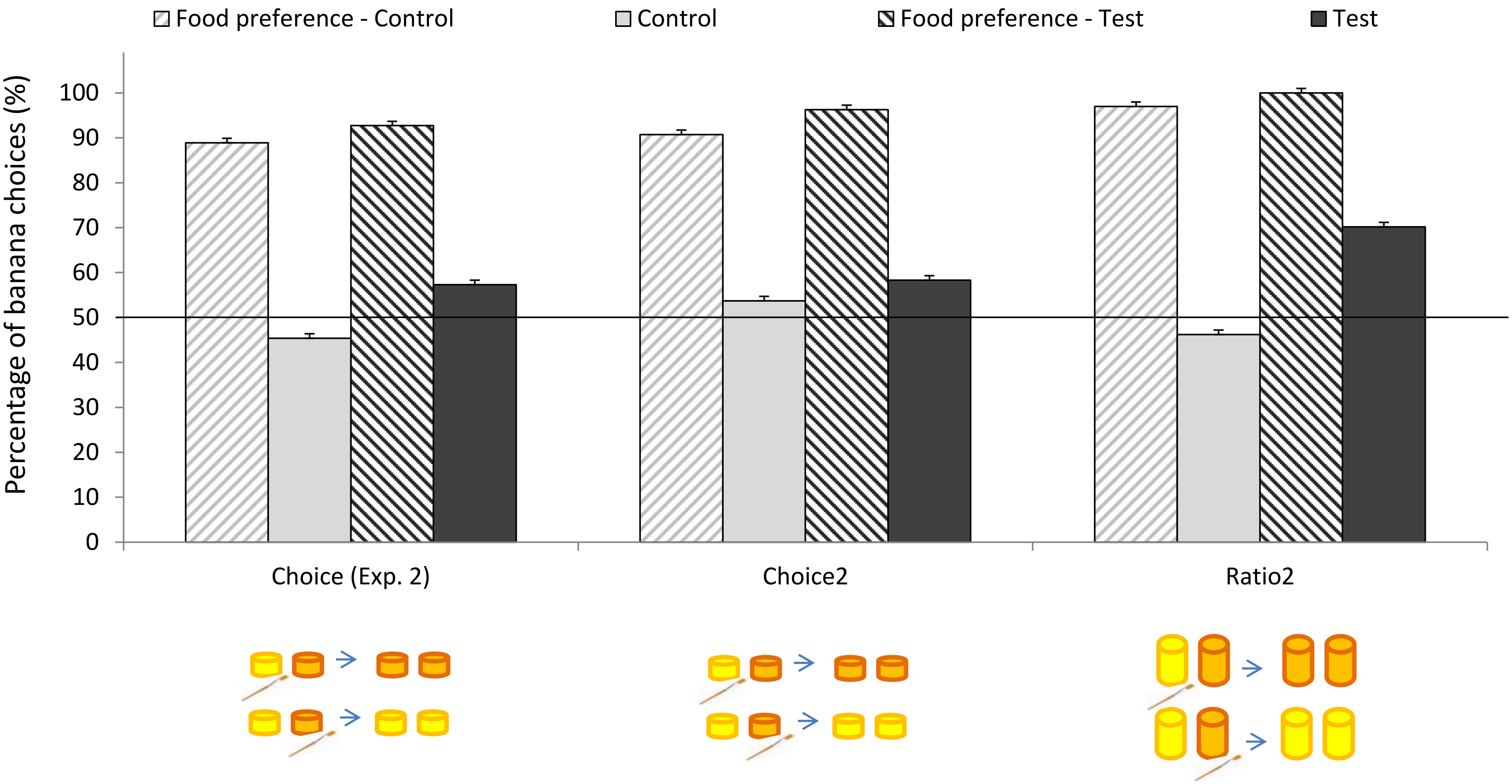
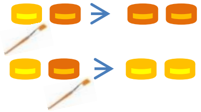
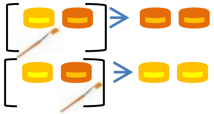
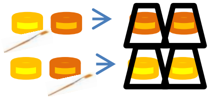
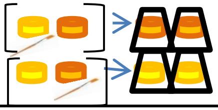
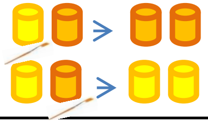
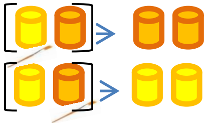

▨ Food preference - Control
☐ Control
▧ Food preference - Test
■ Test
Percentage of banana choices (%)
100
90
80
70
60
50
40
30
20
10
0
Choice (Exp. 2)
Choice2
Ratio2

Table 1

	Main task	**Food preference task**
Experiment 2 – Choice task: **Test condition**	Observe manipulations and choose. Banana becomes carrot (2 trials) and carrot becomes banana (2 trials). ►*Test: B-choice* 	Banana vs. Carrot (4 trials) ►*Test: B-preference* 
Experiment 2 – Choice task: **Control condition**	Choose without observing manipulations. Banana becomes carrot (2 trials) and carrot becomes banana (2 trials). ►*Control: B-choice* 	Banana vs. Carrot (4trials) ►*Control: B-preference* 
Experiment 3 – Inhibition task: **Test condition**	Observe manipulations and choose. Transformed stimuli are occluded. Banana becomes carrot (2 trials) and carrot becomes banana (2 trials). ►*Test: B-choice* 	Banana vs. Carrot (4 trials) ►*Test: B-preference* 
Experiment 3 – Inhibition task: **Control condition**	Choose without observing manipulations. Banana becomes carrot (2 trials) and carrot becomes banana (2 trials). ►*Control: B-choice* 	Banana vs. Carrot (4trials) ►*Control: B-preference* 
Experiment 3 – Memory task: **Test condition**	Observe manipulations and choose. Only one food stimulus is used. Paper becomes banana (4 trials). ►*Test: B-choice* 	Banana vs. Carrot (4 trials) ►*Test: B-preference* 
Experiment 3 – Memory task: **Control condition**	Choose without observing manipulations. Only one food stimulus is used. Paper becomes banana (4 trials). ►*Control: B-choice* 	Banana vs. Carrot (4trials) ►*Control: B-preference* 
Experiment 3 – Ratio task: **Test condition**	Observe manipulations and choose. Larger food stimuli are used. Big banana becomes big carrot (2 trials) and big carrot becomes big banana (2 trials). ►*Test: B-choice* 	Banana vs. Carrot (4 trials) ►*Test: B-preference* 
Experiment 3 – Ratio task: **Control condition**	Choose without observing manipulations. Larger food stimuli are used. Big banana becomes big carrot (2 trials) and big carrot becomes big banana (2 trials). ►*Control: B-choice* 	Banana vs. Carrot (4trials) ►*Control: B-preference*