Capuchin Monkeys Individuate Objects Based on Spatio-temporal and Property/Kind Information: Evidence from Looking and Reaching Measures

Verena Kersken[1], Da Zhang[1], Juan-Carlos Gomez[1], Amanda M. Seed[1*], and Derek Ball[2]

[1]School of Psychology and Neuroscience, University of St Andrews, St Andrews, Scotland
[2]Department of Philosophy, University of St Andrews, St Andrews, Scotland

*Corresponding author (Email: AMS18@st-andrews.ac.uk)

Abstract – A core component of any folk physical understanding of the world is object individuation - the cognitive ability to parse sensory input into discrete objects. Whereas younger human infants use spatio-temporal information to individuate objects, they do not use property and kind information until one year of age. Some researchers propose that object individuation based on property/kind information depends on language acquisition and sortal concepts. However, there is evidence that preverbal infants and nonhuman animals also use both types of information. The present study aimed to further explore the evolutionary origins of object individuation by testing a new-world monkey species, capuchin monkeys (*Sapajus* spp.), using both manual search (*n* = 29) and looking time (*n* = 23) measures. In Spatio-temporal trials, subjects saw one or two objects dropped into a box, but always found (or saw) only one. In Property/kind trials, subjects saw either object A or B being dropped into a box and then always found (or saw) object A. The capuchin monkeys looked longer and searched more on inconsistent trials – with outcome differing in quantity or in kind - which suggested that they had expectations based on both spatio-temporal and property-kind representations. Looking time and search measures gave convergent results at the group but not at the individual level. Our results add to the existing evidence contradicting the linguistic hypothesis of property/kind individuation. However, contrary to recent discussions, we argue that these and related results can be explained without appealing to the notion of sortal concepts or multiple representational systems, and suggest that a full picture of the ontogeny and phylogeny of object individuation requires further empirical and theoretical research.

Keywords – Object individuation, Capuchin monkeys, Property/kind information, Spatio-temporal information

One of the most basic components of the 'folk physics' of adult humans is the ability to represent the environment around them as filled with discrete objects – trees and trucks, coffee cups, and grains of sand. We think of these objects as belonging to kinds (as being trees or coffee cups or grains of sand), as having different properties (as being red or blue, round or square), and as occupying different places and tracing different, typically continuous, paths through space and time. All of these ways of thinking about objects inform our expectations about what we will find in our environment and is the basis for more complex understandings and manipulations of the physical structure of the world. For example, if you leave two pencils in a box, you expect to find the two pencils with the same size, color, and shape later. You would be surprised to find two bigger pencils, or four pencils, or two butterflies. How do non-linguistic creatures like preverbal children or non-human animals, represent objects in the world? Do they rely on the same information as adult humans: namely information about property or kind, and location in

time and space? Or do they use only some of these features to guide their expectations about their environment?

In a classic study, Tinklepaugh (1928) hid a piece of banana in full view of a rhesus macaque. He then surreptitiously switched the piece of banana with lettuce, a much less preferred food. When the monkey uncovered the food, he showed behavior that Tinklepaugh interpreted as "indicative of disappointment" – the monkey continued to search for the food in and around the hiding place, shrieked, and ignored the lettuce. While Tinklepaugh thought this observation was evidence that the monkey represented qualitative aspects about the food, he cautiously concluded "there is no evidence of the true nature of these representative factors" (Tinklepaugh, 1928, p. 236).

Recent work in developmental and comparative psychology has translated Tinklepaugh's observations into more refined paradigms, using looking time and violation of expectation techniques aiming to show the true nature of the representations involved in object individuation. In a particularly influential study, Xu and Carey (1996) presented infants with a display in which objects could be moved behind a screen. In one condition, infants saw one object (a toy duck) emerge from behind the screen, then return, followed by another object (a ball) emerging from the same screen and returning. The screen was then dropped to reveal either just one of the objects, or both objects. Xu and Carey (1996) found that 12-month-old infants use property/kind information to represent objects: showing them two kinds of objects one-at-a time led 12-month-olds to anticipate two objects behind the screen. In contrast, 10-month-olds did not anticipate two objects when they received only property/kind information, although they did when they received spatio-temporal information (i.e., when they saw both objects at the same time). This led the authors to argue that 10-month-olds mainly rely on spatio-temporal information when tracking objects (further supporting a large body of previous research on so-called Spelke-objects), which suggests that the ability to use spatio-temporal information to make inferences about objects is innate as it is present in very young children (e.g., Spelke et al., 1995). Xu and Carey's findings have since been extended to investigate a number of related questions, including questions about exactly which properties and kinds influence object representations (for example, Bonatti et al., 2002) and about the age at which property and/or kind information matters (for example, Wilcox, 1999; Wilcox & Baillargeon, 1998).

Xu and Carey (1996) claimed that their results could be explained in terms of the philosophical notion of *sortal concepts*. Sortal concepts are concepts of kinds of things, typically thought to correspond to common nouns, such as "ball," "computer," or "person," that capture the individuation and persistence conditions for things of that kind. Proponents of sortal concepts claim that they are necessary to answer questions like "How many?" (how many books, how many pages, or how many molecules?), and "Is it the same?" (we may make a new statue from the same lump of clay). And proponents of sortal concepts typically claim that each object falls under exactly one sortal. A statue may be constituted by a lump of clay, but (on the sortal view) the lump and the statue are distinct objects because they fall under distinct sortals and thereby have distinct persistence conditions; for example, the lump can survive the destruction of the statue.

Two Systems for Object Individuation?

Xu and Carey (1996) suggest that the change in infants' performance between the ages of 10 and 12 months can be seen as strong evidence for a fundamental change in the way infants conceive of objects, namely that 12-month-olds possess sortal concepts that 10-month-olds do not: whereas 12-month-olds possess a variety of sortals like "ball" and "duck" (and hence are capable of representing that a ball is not a duck), 10-month-olds would possess only a single sortal – roughly glossed as "bounded physical object" - which is not sufficient to individuate the two objects in the task.

Carey and Xu (2001) argue further that these data can be explained on the view that infants' representations of objects are the result of a specialized, encapsulated system, which indexes and tracks bounded physical objects (so called "Spelke objects;" but see Pylyshyn, 2001, for an alternative explanation based on perceptual object tracking with no conceptual content). Information about these objects is stored in symbolic representations called *object files* (Carey & Xu, 2001). Carey and Xu (2001)

therefore suggest that there are at least two systems for individuating objects: the object-file system, which largely makes use of spatio-temporal information and shows clear signature limits (which would underlie sortal concepts related to Spelke objects), and the later-developing kind-based system (which would underlie other sortal concepts, such as "ball" and "duck").

Xu (2002, 2007) suggests further that language acquisition, namely the ability to understand and produce sortal nouns, plays a crucial role in the development of the kind-based system. Xu and Carey (1996) proposed that the alleged developmental shift between 10 and 12 months of age in how objects are represented is associated with the acquisition of sortal nouns. In short, Xu and Carey defend the following hypothesis:

Linguistic Sortal Concept Hypothesis
(A) *Sortal Gain:* 12-month-olds are sensitive to property/kind information that 10-month-olds are not because 12-month-olds possess sortal concepts that 10-month-olds do not.
(B) *Sortal-Language Link:* 12-month-olds' possession of sortal concepts is closely connected with their linguistic abilities, specifically the ability to use sortal nouns.

Empirical Challenges to the Two-System Account

However, there is evidence that Xu and Carey's bold hypothesis was premature. Linguistic Sortal Concept Hypothesis suggests that non-linguistic (or partially linguistic) beings, like young children and nonhuman animals, would perform differently in Xu and Carey's (1996) task. However, other studies show sensitivity to property/kind information in much younger infants. For example, Wilcox and Baillargeon (1998) found that even 7.5-month olds' expectations are influenced by information about object properties in a simplified version of Xu and Carey's procedure (see also Wilcox & Chapa, 2002; Wilcox & Schweinle, 2002).

Moreover, there is also evidence that nonhuman animals are sensitive to property/kind information. For example, Uller et al. (1997) using a looking-time procedure very much like Xu and Carey's, found that rhesus macaques were sensitive to property/kind information (as well as spatio-temporal information). Later studies with free-ranging rhesus macaques (Phillips & Santos, 2007; Santos et al., 2002) came to similar conclusions, measuring the monkeys' "searching time," which included both looking and reaching into or around a box after watching the experimenter bait it with food items but finding only the non-matching pre-baited food. The monkeys engaged in significantly longer searches when the food found did not match the baited food, suggesting they represented the kind of food.

Studies with great apes have provided additional evidence that some nonhuman animals represent property/kind information. Mendes et al. (2008) found that great apes (bonobos, chimpanzees, and gorillas) are sensitive to both spatio-temporal and property/kind information, searching more inside a box if the object or amounts they initially found were different to what they saw being placed. This finding was replicated in follow-up studies showing that great apes are sensitive to property/kind information, with either searching or begging behavior measures (e.g., Bräuer & Call, 2011; Mendes et al., 2011).

These comparative studies not only challenge the Linguistic Sortal Concept Hypothesis; they also cast doubt on Xu and Carey's (Carey & Xu, 2001; Xu & Carey, 1996) two representational system account. The two-system account was motivated by the empirical evidence of a dissociation between individuation by spatio-temporal information and by property/kind information in human infants: 10-month-olds seem to do the former but not the latter. If the two-system account is correct, we would expect to find species that (like 10-month-olds) have only the object-files system and hence individuate objects only using spatio-temporal information, especially if one assumes that property/kind individuation is closely linked to recent evolutionary developments (if not the ability to use language, which the comparative evidence to date would throw into question, perhaps some precursor to this, such as abstract representations of objects and their properties). However, there is currently no evidence of dissociation between the two ways of individuating objects in nonhuman animals. All the nonhuman primate species tested so far showed competence on both spatio-temporal and property/kind object individuation tasks (Bräuer & Call, 2011; Mendes et al., 2008; Santos et al., 2002; Uller et al., 1997). There is even evidence

that domestic dogs and chicks succeed in similar tasks (Bräuer & Call, 2011; Fontanari et al., 2011, 2014). The proponent of the two-system account might argue that these nonhuman animal subjects are not truly individuating objects by property or kind information, but are simply relying on featural tracking (e.g., looking for some "yellowness" instead of the banana slice). However, further evidence revealed that apes' and macaques' searching behavior is shaped by kind information more than by superficial features such as color and shape (Cacchione et al., 2016; Phillips et al., 2010). For example, Cacchione et al. (2016) found that apes searched longer when they saw an orange-colored banana go in and found a carrot, than when they found the orange dyed banana or even an un-dyed (yellow) banana (i.e., a superficially different object of the same kind). This evidence suggests that nonhuman primates did not just rely on featural tracking but may engage in sortal object individuation (Cacchione et al., 2016; Cacchione & Rakoczy, 2017; Phillips et al., 2010; Rakoczy & Cacchione, 2019).

Taken together, findings from comparative research do not favor the two-system account. Xu and Carey's idea of a late-developing, linguistically-based, second system for kind-based representation in human infants would predict exclusively spatio-temporal individuation in nonhuman animals; but, there is currently no evidence in phylogeny that a species can use spatio-temporal information to individuate objects but fail to individuate objects using property/kind information.

However, the fact that we currently lack evidence for the two-systems account does not show that the account is false. There are three possibilities:

(1) There is only one system for object individuation. The developmental change between 10-month-olds and 12-month-olds must be explained in some other way.

(2) There are two systems for object individuation, but the ability to individuate objects using property and kind information is evolutionarily older than the Linguistic Sortal Concept Hypothesis suggests. We have not found the species, or the key point in evolution, corresponding to that developmental change between 10-month-olds and 12-month-olds.

(3) There are two systems of object individuation, but they evolved simultaneously in phylogeny and we will never find a species with only one of them. The existence of the two systems is only temporarily visible during ontogeny.

A full examination of possibilities (2) and (3) requires a more detailed picture of the evolutionary history of the ability to individuate objects using property/kind information. As discussed above, some primates (apes and macaques) have shown evidence for the ability to individuate objects using property/kind information, and there is some evidence that dogs and chickens may show some sensitivity to property/kind information in comparable paradigms. This may mean that this ability evolved in a common ancestor of primates, dogs, and chickens and this is why it is shared despite other significant cognitive differences between these species. Alternatively, the ability may have evolved separately but convergently in species tested to date. To address this issue, we need to investigate object individuation skills in further species. One aim of the present study is to probe, for the first time, the object individuation skills of a New World monkey species, capuchin monkeys.

Two Ways of Measuring Object Individuation

A second, related aim of our study is to address the methodological and conceptual issue of looking time vs action-based measurements of violation of expectation in object representation studies. The majority of studies in the literature on object representation in infants and nonhuman primates focus on two measures of violation of expectation: looking time or reaching time (or number of reaches). It is controversial in some research domains whether these two measures are tracking the same ability, since sometimes the performance revealed by different measures does not converge. For example, in several solidity tasks when subjects are presented with a pair of events in which an object either stops at or moves through a solid barrier, human infants usually pass the looking task (Spelke et al., 1992), while older toddlers may fail the searching version of the task (Berthier et al., 2000; Hood, Carey, et al., 2000). This kind of dissociation was also found in adult monkeys who had adequate performance ability for action (Cacchione & Burkart, 2012; Hauser, 2001; Santos & Hauser, 2002). Moreover, the performance of the

same population measured by looking and searching tasks may not correlate. For example, in an invisible object displacement task, younger toddlers passed the looking time version, but failed the searching version (Hood et al., 2003). This evidence challenges the interpretation that success on tasks using looking measures reveals full-fledged representational ability. Two alternative ways of understanding these data have been proposed. The first interpretation suggests that searching requires "stronger" representations than looking (Munakata, 2001). In other words, looking tasks may require only detecting a violation (post-hoc or "postdiction") but without necessarily understanding the violation and making accurate predictions, which would be required to guide the action (Cacchione & Burkart, 2012; Hood et al., 2003). Dissociation between looking and searching tasks may reveal multiple systems, such as one for "what" or object recognizing, and another for "where" or object-directed grasping (Goodale & Milner, 1992; James et al., 2003; Leslie et al., 1998; Milner & Goodale, 2008). However, a second explanation is that the dissociation between measures does not reflect the existence of different systems, but simply that the looking task is more sensitive than the reaching one, because it is less likely to be confounded with other skills needed to successfully retrieve an object over multiple trials. For example, 2.5-year-old toddlers reach repeatedly to one location, perhaps revealing that the limiting factor is the immature executive functions of toddlers, making it difficult to resolve the conflict created by predictions based on object tracking and past reinforcement (Berthier et al., 2000).

We have no definitive view on how best to explain divergence between looking and reaching time measures; different explanations might be correct in different cases. We nonetheless maintain that a detailed study of object individuation in a given species requires evidence from both kinds of measure. Relying on looking measures may result in attributing a species a full competence that it in fact lacks, while relying on reaching measures fails to detect a weaker form of representation or competence, which might be masked due to confounding factors. Most importantly, differential patterns of dissociation/convergence between the two sources of information (spatio-temporal and property/kind) could reflect the presence of multiple systems for object individuation.

Considering these debates, it is striking that there are no studies comparing these different measures in nonhuman animals on the same object individuation task: Uller et al. (1997) measured looking time in macaques, whereas Mendes et al. (2008) measured reaching time in apes, but the results cannot be directly compared because the experiments involve different species and different experimental designs. The "searching time" measure of Santos and colleagues (Phillips & Santos, 2007; Santos et al., 2002) is not useful for these purposes because it combines (and so does not distinguish) looking and manual search. By contrast, there is evidence that measures of visual search and manual action converge in human infants, in terms of the age of emergence. Van de Walle et al. (2000) replicated Xu and Carey's (1996) results using a manual search paradigm: they found that 12-month-olds but not 10-month-olds reach longer into a box when their expectations about property and kind are violated. However, Wilcox and Baillargeon's (1998) finding of competence in younger infants using a simpler paradigm has to date been conducted only with looking time as the dependent variable. Reaching measures are difficult to conduct in this population.

In summary, the present study has two main aims. First, in order to provide a fuller picture of the evolutionary history of object individuation, we tested a new species, capuchin monkeys with an adaptation of the box task used in previous primate studies to investigate their property/kind and spatio-temporal individuation skills (Mendes et al., 2008; Santos et al., 2002). This is the first New World monkey species to be tested with this paradigm. Capuchin monkeys are of special interest as they are a species of new world monkey that would benefit from sophisticated object representations as a large-brained extractive forager. This species, though more distant to human beings in phylogeny, has shown many cognitive abilities that are similar to apes' (for example, with respect to causality and tool use, see: Fujita et al., 2003; Penn & Povinelli, 2007; Povinelli, 2000), but has not previously been studied in this object individuation context.

We did not have clear predictions about how capuchin monkeys would perform. They might show competence in both individuation by spatio-temporal information and property/kind information, as the other nonhuman species tested to date. Or, if property/kind object individuation is cognitively

different to spatio-temporal individuation, we might find dissociation in this species; for example, capuchins might individuate objects by spatio-temporal information only.

Secondly, we measured both looking and searching behaviors with exactly the same apparatus in the same subjects. The simple, but important, issue of direct comparison of the different measurements has not been addressed before in nonhuman primates. We aimed to explore the relationship between the two measures on both group level (whether the monkeys pass or fail tasks with both measures) and the individual level (whether the monkeys' performance on either measure correlated). If spatio-temporal and property/kind individuation rely on different systems or cognitive abilities, we might find differential dissociations in each type of tasks (for example, convergence between looking and reaching may occur in the Spatio-temporal task, but not the Property/kind task). Parallel dissociation across both types of individuation would also make a strong case for different systems (for example if both measures detect competence in the Spatio-temporal condition but not the Property/kind condition). While convergence in both conditions using both measures would be more consistent with a single system or similar cognitive abilities.

Method

Participants and Site

In total, 29 Brown Capuchin monkeys (*Sapajus sp.*) from two social groups participated in the present study. Among them, 25 individual monkeys (14 male) completed trials in the looking time task, and 29 individuals (15 male) completed the manual search trials (for details of participation and completion, please see Table 1). The monkeys' ages ranged from 2 to 21 years ($M = 8.18$, $SD = 5.01$).

Table 1

Total Number of Valid and Excluded Trials for Each Condition in Each Paradigm

| Measure	Condition	No. of trials for analysis	No. of excluded trials
Looking Time	PKC	20	2 (1 error, 1 no reaction)
	PKI	23	0
	STC	18	3 (1 error, 2 no reaction)
	STI	21	2 (1 error, 1 video material)
Manual Search	PKC	19	7 (2 error, 4 no reaction, 1 video material)
	PKI	27	2 (1 error, 1 no reaction)
	STC	25	3 (3 no reaction)
	STI	23	4 (1 error, 3 no reaction)

All but one monkey was born in captivity and raised by their mothers in a social group. One monkey was wild-born and rescued from the pet trade with an unclear rearing history. All monkeys were housed at the Living Links Research Centre at RZSS Edinburgh, Scotland. Monkeys in the Living Links Centre were housed in two equally sized groups (East and West) with identical enclosures and provisions. Both Capuchin groups lived in a mixed-species community with squirrel monkeys (*Saimiri sciureus*). All monkeys had previously taken part in behavioral studies at the site. Indoor and outdoor enclosures were furnished with climbing frames, vegetation, and visual barriers. Monkeys could move freely between the indoor and outdoor enclosure via doors and tunnels. The research cubicles were situated in a separate area between the indoor and outdoor enclosures. When no research was taking place, the monkeys could use the cubicles as an additional access route to the outdoor enclosure. More details on Living Links, primate housing, and husbandry can be found in MacDonald and Whiten (2011). The monkeys were not food or water deprived during the study. Participation in the study was voluntary. Monkeys received supplemental food rewards (sunflower seeds, dates, and grapes) for participating. The study was granted ethical approval by the University of St Andrews' Animal Welfare and Ethics Committee.

Apparatus

The apparatus consisted of an opaque plastic box (35 cm x 25 cm x 40 cm) with a square opening at the top through which the experimenter could drop food rewards into the box. The box was placed on a projector table with wheels so it could be easily moved around. The main compartment floor of the box was padded with carpet to absorb any noise of falling objects. The box had a double floor that was invisible from the outside and prevented rewards that were dropped into the box from reaching the bottom compartment (see Figure 1). This allowed the experimenter to store food rewards without the subject's knowledge. The double floor could be removed, and we used this feature in the familiarization phase to allow monkeys to see a food item drop from the experimenter's hand and fall into the main compartment of the box. The lower fifth of the box was open and could be covered by different Plexiglas and plastic sliders that varied for each experimental paradigm.

Figure 1

CAD Drawing of the Apparatus from the Experimenters' Point of View
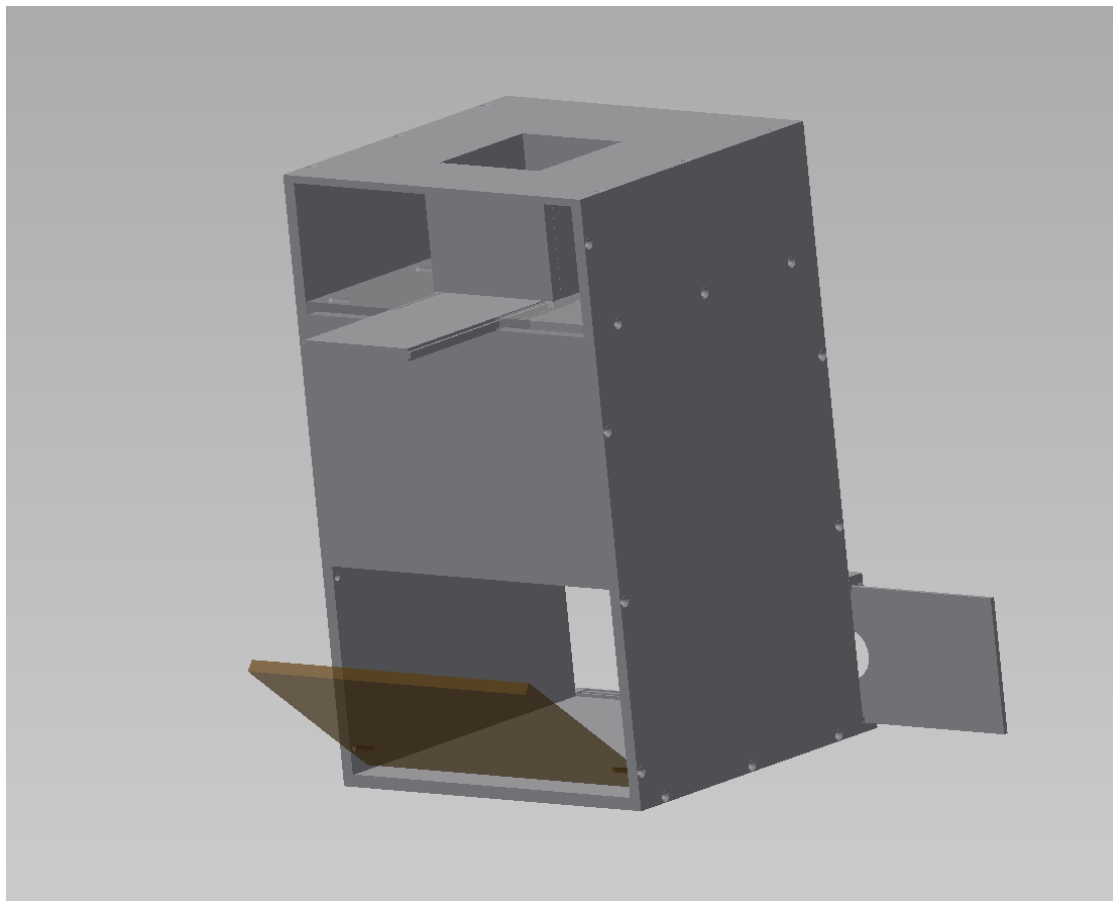
Opaque Slides

For both trials (either familiarization or experimental trials under either reaching trials or looking condition), an opaque (black) Plexiglas sliding door was used to cover the lower front opening of the box to block any visual or manual access into the box. This slide was placed on the top of other slides (in the track to the monkey's side).

Looking Slides

For the looking time part of the study, a transparent Plexiglas sliding door was used. By removing or inserting the opaque plastic slider, the experimenter manipulated the monkey's visual access to the contents of the box. Additionally, we placed an additional LED light inside to guarantee the visibility of the content.

Reaching Slides

For the manual search conditions, an opaque slider with an elongated reach hole (about 12 cm x 3 cm, rounded rectangular) was locked into the lower opening of the box. The reach hole lined up with the holes in the cubicles the monkeys watched the study from. By removing or inserting the opaque plastic slider, the experimenter manipulated the monkey's manual access to the contents of the box. At the back of the box, facing the experimenter, was a flap that allowed the experimenter manual access to the main compartment of the box. This was used to surreptitiously place rewards in the box.

GoPro cameras (models Hero 4 silver and Hero 4 session) were used to record the experiment. One camera was placed next to the apparatus on a Gorillapod, roughly at the level of the subject's face, and recorded the monkey's face during the experiment. The camera filmed the monkey and the surrounding cubicle. We used footage from this camera to code the looking time and frequency to the content of the box. The second camera was placed on a shelf behind the experimenter and recorded the entire scene. We used this footage to check for any procedural errors or factors in the immediate environment that would invalidate a trial. For the manual search trials, an additional camera was placed inside the box, filming the inside of the box through the transparent double floor. This footage was later used to code for reaching behavior inside the box. Cameras were controlled using a GoPro remote.

Food rewards consisted of sunflower seeds, dates, and grapes. Monkeys were rewarded for coming into the cubicle and received an additional reward at the end of each trial. The type and amount of food rewards was regulated by keepers at Living Links.

Procedure

Testing Procedure

Each monkey could undertake one trial per session, a maximum of two sessions a day (one between 11:15-12:45 AM and one between 14:15-15:45 PM). During each session, a monkey could visit the research cubicle though the entrance to their enclosure on a voluntary basis. Once they entered, the experimenter (V. K.) shut the sliding door between cubicles to separate the subject from its group. Then the experimenter started the trial in front of the monkey (about 50 cm from the cubicle window). The monkey was returned to the group by opening the slide when the trial finished, or if it signaled a lack of motivation to continue by touching the slide or otherwise disengaging with the task.

Food Preference

We conducted a food preference test with all participating monkeys to find two food rewards of equal value to the monkeys. Monkeys were presented with two different food rewards (dried mango, dried apple rings, grapes, dates, shelled peanuts) in the experimenter's palm and were allowed to choose one of them. We repeated this test several times. We found that the monkeys equally preferred grapes and dates, and these food types were visually sufficiently different to be used in the study.

Familiarization

Before the experimental trials, all subjects went through a familiarization phase that consisted of several trials. The monkeys in the East and West group were presented slightly different familiarization trials, since they started on the looking time trials (West) or manual search task (East). All monkeys completed the familiarization trials only once. We did not conduct a new familiarization phase when the monkeys switched paradigms to avoid habituation.

The differences between familiarization trials for the looking-first group and searching-first group were that (1) the monkeys in the searching-first group were allowed to reach for the food after the demonstration, whereas the monkeys in the looking-first group could only watch before getting the food from the experimenter after roughly 15 s; (2) an object retrieval trial was provided for monkeys in the searching-first group only. Detailed procedures of familiarization tasks were as below.

0) Object retrieval (only for searching-first group): Subjects were presented with the apparatus and a food reward in the bottom compartment. Monkeys were then given the chance to manually retrieve the food from the box via the access hole. Only when monkeys completed this trial were they presented with the other familiarization trials.
1) Transparent trials: The monkeys had full visual access to the main compartment of the box. The double floor in the box was removed. The experimenter placed a food reward next to the box so the monkey could see it. Monkeys watched the experimenter drop the food reward into the box and fall into the main compartment. Then the monkeys were given the chance to manually retrieve the food from the box via the access hole or get the food from the experimenter.
2) Opaque trials: Prior to the trial, the experimenter placed a food reward into the main compartment of the box. The monkey did not see this action. The opaque slider was inserted so that the monkey could not see into the main compartment of the box. The double floor was inserted into the box without the monkeys' knowledge. The experimenter placed a food reward next to the box and then dropped the food rewards into the box in the full view of the monkeys. The opaque slider was then removed to allow visual or manual access to the main compartment of the box. Monkeys could get the food themselves or, if they found the food but had difficulty retrieving it, from the experimenter. The whole procedure aimed to give the monkeys the illusion that the food fell straight to the bottom while, in fact, the food was stopped by the hidden floor.

For all familiarization and experimental trials, the experimenter hit the bottom of the box with their free hand when dropping the reward to mask any possible audible cues to the reward dropping either into the main compartment or onto the double floor.

Experimental Trials

The procedure, adapted from Santos et al. (2002) and Mendes et al. (2008), included two tasks (Spatio-temporal task and Property/kind task) with two versions each, measuring either looking or manual searching behavior. Each task had a consistent condition and an inconsistent condition. Thus, there were four conditions (for key differences among conditions, see Table 2) for each measure and each subject received only one trial per condition. We counterbalanced across the two study groups (namely the Searching-first and the Looking-first group, or the West and East Group) whether the manual search or looking time trials were presented first. Within each block of each measure, the four kinds of experimental trials were presented in random order.

Table 2

Baited and Retrieved Food Items for Each Condition

Condition	Food inserted	Food retrieved	"Missing" food
Property/kind Consistent (PKC)	A	A	-
Property/kind Inconsistent (PKI)	B	A	B
Spatio-temporal Consistent (STC)	A	A	-
Spatio-temporal Inconsistent (STI)	AA	A	A

Note. A and B denote different kinds of food rewards (grapes or dates). The "missing" food indicated that the food item should be found and will be given to the monkey by the end of the trial as if it was stuck somewhere but still in the box. Note that the retrieved food was always the same across trials.

All experimental trials followed the same general procedure except for the number or kind of rewards that were inserted into the box (the input to the box) or that were retrieved from the box (i.e., the "missing" object) by the experimenter after the recording periods. Specifically, what a monkey could see or find (the output of the box) at the end of each trial was always the same across conditions (Table 2). Therefore, any performance difference found could not be explained by different outcomes, but only the input and the subjects' expectation based on the input.

Before each trial, the experimenter preloaded the main compartment of the box with a food reward and covered the front opening with first a looking slide or a reaching slide, and then an opaque slide. After the preparation, the monkey was invited into the research cubicle and given a small reward (sunflower seed) for participating. The experimenter showed the target food reward(s) (date or grape) to the monkeys, lifted the reward(s), and, in her closed hand, dropped it/them into the box (onto the double floor). Note that if there were two pieces of food rewards in this trial, the experimenter showed, lifted, and dropped the food one piece at a time. Whilst dropping the food, with her other hand, she hit the bottom of the box, to mask any dropping sounds from the food reward. The opaque slide was then removed, and the monkey was either allowed 25 s of visual, or 35 s of manual access to the contents of the box. After these periods, the experimenter inserted the opaque slide back, removed the box by pulling back the trolley table, and opened the back flap to inspect the contents. In the looking time trials, monkeys were now given the food they had seen. In the inconsistent conditions, after the monkeys were either handed the reward they had seen, or when they no longer had manual access to the box, the experimenter opened the back flap, rummaged round the box to retrieve an additional food item, namely the "missing" object, as if it had been "stuck," and then handed it to the monkey, marking the end of the trial. This step was to maintain the illusion that this was a "normal" box in which inserted food fell straight to the bottom.

Predictions

If monkeys can individuate objects on the basis of spatio-temporal information, we would expect them to look longer, and search longer, in the box in the inconsistent condition than in the consistent condition, because they would expect a second object to be present. Similarly, if monkeys are able to individuate objects on the basis of property/kind information, we would expect them to look longer, and search longer, in the box in the inconsistent condition than in the consistent condition, because they would expect a different kind of object in the box. Because our hypotheses were directed, we used one-tailed tests in the statistical analyses.

Coding

We coded our videos using the program ELAN 4.9.2 (http://tla.mpi.nl/tools/tla-tools/elan/; Lausberg & Sloetjes, 2009) on a MacBook Pro computer. Prior to coding the looking time, the videos were cropped to the duration of the specified coding period (25/35 s) and renamed. This eliminated a potential coding bias as all information about the experimental conditions were removed prior to coding. We coded the monkeys' looking time at the contents of the box for 25 s after the content was revealed to

them as indicated by the removal of the opaque slider. The main dependent measure was the total looking time at the target (contents of the box), which was defined as looks that are directed at the target area of the box (lower 25%). The looking direction was discerned from the monkeys' face orientation. We coded the monkeys as looking at the target when their face was oriented in a way that allowed them visual access to the contents of the box (e.g., face opposite opening in the box, face lowered or angled to allow viewing the target). We coded each separate instance of a look within the duration of the coding period. The duration of a looking bout was measured with an accuracy of 2 frames (40 ms). We measured the overall time spent looking at the target (looking time) and the total number of looks (looking bouts) within the coding period.

In the manual search condition, we coded the monkeys' reaches into the box. The coding period started as soon as the opaque slider was removed from the access hole, allowing the monkey to reach into the box. Reaches were operationally defined as the insertion of any part of the hand into the box, as seen by the internal camera. We calculated the total duration of reaches into the box in the specified coding period (reaching time), and the total number of individual reaches into box in the coding period (reaching bouts).

In addition to manual reaching, we also coded when the monkeys peered into the box, in case sometimes a monkey noticed (and was surprised by) the violation but did not want to reach in. We introduced this coding category as we noted that, in general, the monkeys were somewhat shy about inserting their hand into the box, but went through considerable effort to gain visual access to the contents of the box (pressing their face to the access hole, bending down to look into the box). We coded 'peering' as a monkey positioning their head and body in a way that would allow them to look into the box through the access hole. This included crouching down in front of the access hole and pressing face and forehead towards the access hole to look into the box. We coded the total time the monkeys spent peering into the box (peering time) as well as the number of individual peers (peer bouts). Peer bouts were separated by the monkey lifting their head or moving their body in a way that would interrupt their visual access to the contents of the box.

Data Analysis

The data was analyzed using the statistical package SPSS 23.0. As the data was largely non-normally-distributed (see Table 3 for the results of the normality tests), we conducted Wilcoxon Signed Ranks tests instead of ANOVA to compare the median looking, reaching, and peering time across the conditions. We used the same test to compare the median looking, reaching, and peering bouts. We conducted additional Mann-Whitney-U tests to account for the group the monkeys belonged to, i.e., whether they were first presented looking time or manual reaching trials. Moreover, we analyzed the correlation between measures using difference scores, which were calculated from the performance difference between Consistent versus Inconsistent trials of each task (Property/kind vs. Spatio-temporal) for each measure.

To assess interrater reliability, a naïve coder scored a random sample of 20% of the trials in the looking time and manual search conditions. Interrater reliability was high for both conditions (Looking time: Cohen's kappa; $r = .92$; Manual search and peering: Cohen's kappa, $r = .88$).

Results

A number of trials had to be excluded from the analysis because of experimental errors (food reward rolling out of sight or reach, disruption in the environment), the video material not being suitable for coding (monkey's face out of sight, strong glare) or because the monkeys did not react or left the experiment early. Table 4 illustrates the total number of analyzed and excluded trials for each condition.

Table 3

Shapiro-Wilk Normality Test

Measures	Dependent Variable	Condition	Statistic	*df*	Sig.
Looking	Looking time	PKC	.863	13	.043*
		PKI	.936	13	.409
		STC	.866	13	.047*
		STI	.88	13	.072+
	Looking bouts	PKC	.836	15	.011*
		PKI	.892	15	.071+
		STC	.87	15	.034*
		STI	.885	15	.057+
	Reaching time	PKC	.925	14	.260
		PKI	.909	14	.153
		STC	.951	14	.574
		STI	.935	14	.358
Searching	Reaching bouts	PKC	.428	14	< .001***
		PKI	.639	14	< .001***
		STC	.297	14	< .001***
		STI	.616	14	< .001***
	Peering time	PKC	.753	14	.001**
		PKI	.889	14	.079+
		STC	.954	14	.626
		STI	.894	14	.092+
	Peering bouts	PKC	.809	13	.009***
		PKI	.832	13	.017*
		STC	.756	13	.002**
		STI	.901	13	.139

Note. Significance indicates non-normality. * $p < .05$; ** $p < .01$; *** $p < .001$; † $p < .10$.

Table 4

Demographics and Participation Information of Capuchin Subjects

	Name	Sex	Age (yr)	Looking time				Manual searching			
				PKC	PKI	STC	STI	PKC	PKI	STC	STI
	Alba	F	5	1	1	1	1	1	1	1	1
	Bear	M	3	1	1	1	1	1	1	1	1
	Diego	M	14	1	1	1	1	1	1	1	1
	Fudge	M	3	1	1	1	1	1	1	1	1
	Hazel	F	3	1	1	1	1	1	1	1	1
	Inti	M	7	1	1	1	1	1	1	1	1
	Lana	F	21	1	1	1	1	1	1	1	1
West group	Luna	F	5	1	1	1	1	1	1	1	1
(Looking-first)	Mekoe	M	8	1	1	1	1	1	1	1	1
	Pedra	F	8	0	1	0	1	0	0	0	0
	Pixie	F	3	1	1	1	1	1	1	1	1
	Rufo	M	7	1	1	1	1	1	1	1	1
	Sylvie	F	13	1	1	1	1	1	1	1	1
	Toka	M	12	1	1	0	1	1	1	1	1
	Torres	M	5	1	1	0	0	1	1	1	1
	Ximo	M	6	1	1	1	0	1	1	1	1
West dropout		-	-	1	0	3	2	1	1	1	1
West summary			7.82±4.89	15	16	13	14	15	15	15	15
	Agnes	F	2	0	0	0	0	1	1	1	1
	Anita	F	19	1	1	1	1	1	1	1	1
	Carlos	M	10	1	1	1	1	1	1	1	1
	Chico	M	7	1	0	1	1	1	1	1	1
	Flojo	M	5	0	1	1	1	1	1	1	1
	Gus	M	2	0	0	0	0	0	1	1	1
East group	Junon	F	16	1	0	1	1	1	1	1	1
(Searching first)	Kato	M	11	1	1	1	1	1	1	1	1
	Lindo	F	5	1	1	0	1	0	1	1	1
	Nena	F	4	0	0	0	0	1	1	0	0
	Penelope	F	11	0	0	0	0	0	1	1	0
	Reuben	M	6	1	1	1	1	1	1	1	1
	Rosa	F	6	0	0	0	0	1	1	1	1
	Willow	F	4	0	1	1	1	1	1	1	1
East dropout		-	-	7	7	6	5	3	0	1	2
East Summary			8.20±5.25	7	7	8	9	11	14	13	12
Overall			8.18±5.01	22	23	21	23	26	29	28	27

Note. The Age column includes the mean±standard deviation of ages in each group. The Condition column reports the number of trials subject(s) participated. STC means Spatio-temporal Consistent; PKC means Property/kind Consistent; STI means Spatio-temporal Inconsistent; PKI means Property/kind Inconsistent.

Looking Time

We calculated monkeys' mean looking time and number of looks directed at the target for each condition (Figures 2 & 3). Monkeys looked longer at the inconsistent events than at the consistent events in both Property/kind, and Spatio-temporal conditions (Property/kind: Wilcoxon test: $z = 2.203$, $n = 20$, $p = .014$, $r = .493$; Spatio-temporal: Wilcoxon test: $z = 1.618$, $n = 17$, $p = .047$, $r = .408$).

We found no significant difference in looking time in each condition between the Searching-first and Looking-first group in the conditions (Mann-Whitney-U: $U(PKC) = 76.5$, $p = .294$, $n = 23$, $r = .222$); $U(PKI) = 38.5$, $p = .336$, $n = 21$, $r = -.213$; $U(STC) = 59$, $p = .238$, $n = 19$, $r = .284$; and $U(STI) = 45$, $p = .393$, $n = 22$, $r = -.192$).

There was no significant difference in the number of looking bouts in the inconsistent and consistent events in both Property/kind and Spatio-temporal conditions (Property/kind: Wilcoxon test: $z = -1.553$, $n = 21$, $p = .060$, $r = .339$; Spatio-temporal: Wilcoxon test: $z = 0.027$, $n = 17$, $p = .489$, $r = -.007$).

There was no significant difference in the number of looking bouts between the Searching-first and the Looking-first group in the conditions (Mann-Whitney-U: $U(PKI) = 52$, $p > .999$, $n = 21$, $r < .001$; $U(STI) = 74$, $p = .324$, $n = 22$, $r = .237$). There was a significant difference between the groups in the condition $U(PKC) = 93$, $p = .034$, $n = 23$, $r = .470$, and a marginal significant difference between the groups in the condition $U(STC) = 68$, $p = .051$, $n = 19$, $r = .468$.
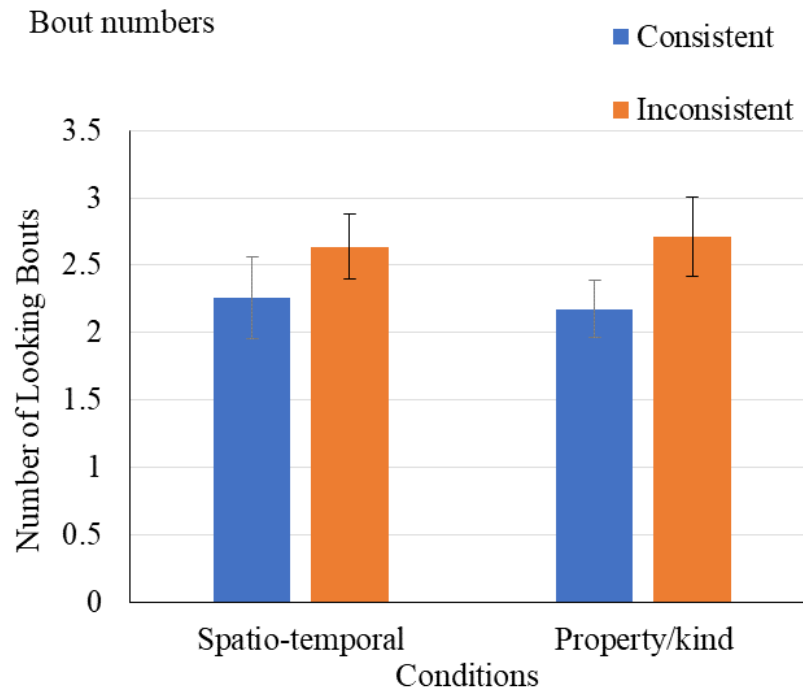
Figure 2

Mean Looking Time (Seconds) at the Target for Each Experimental Condition
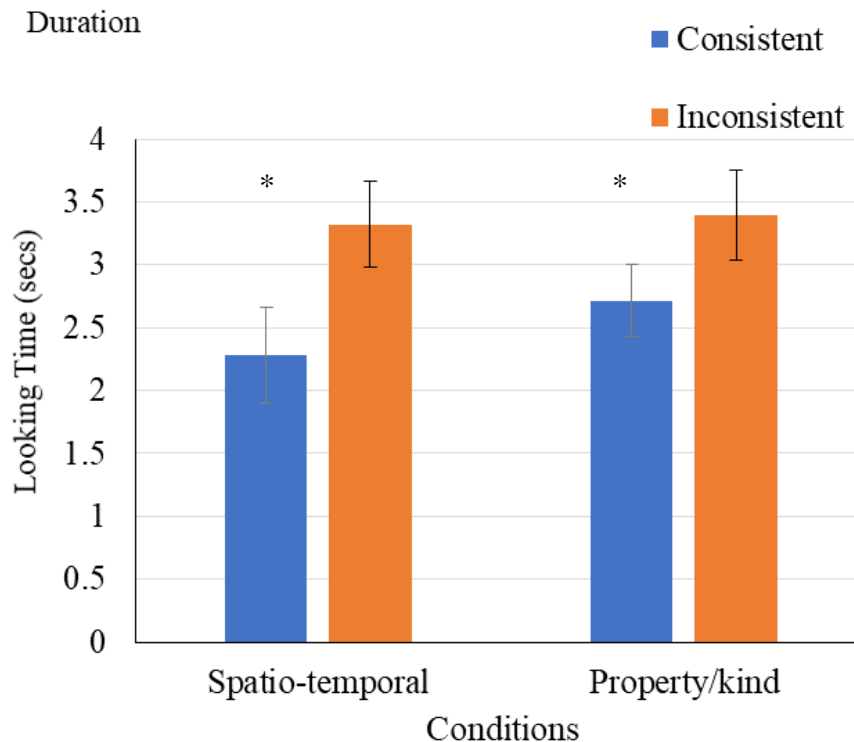
Figure 3

Mean Number of Looking Bouts for Each Experimental Condition

Manual Search

We calculated monkeys' mean reaching time and number of reaches into the box for each condition (Figures 4 & 5).

All monkeys reached into the box at least once to retrieve the food reward. Monkeys reached significantly longer into the box in the inconsistent conditions (Property/kind and Spatio-temporal) than in the consistent conditions (Property/kind: Wilcoxon test: $z = 1.982$, $n = 18$, $p = .024$, $r = .467$; Spatio-temporal: $z = 3.702$, $n = 21$, $p < .001$, $r = .808$).

There were no significant differences in reaching time between the Searching-first and the Looking-first group in all conditions (Mann-Whitney-U: $U(PKC) = 23$, $p = .179$, $n = 19$, $r = -.322$; $U(PKI) = 87$, $p = .920$, $n = 26$, $r = .025$; $U(STC) = 66$, $p = .572$, $n = 25$, $r = -.120$; $U(STI) = 40$, $p = .131$, $n = 23$, $r = -.323$).

Monkeys reached more often into the box during both inconsistent conditions (Property/kind and Spatio-temporal) compared to the consistent conditions (Property/kind: Wilcoxon test: $z = 2.236$, $n = 18$, $p = .012$, $r = .527$; Spatio-temporal: $z = -2.446$, $n = 21$, $p = .007$, $r = .534$).

There were no significant differences in reaching bouts performed by the Searching-first and the Looking-first group in all conditions (Mann-Whitney-U: $U(PKC) = 33$, $p = .639$, $n = 19$, $r = -.227$; $U(PKI) = 61.5$, $p = .243$, $n = 26$, $r = -.287$; $U(STC) = 71.5$, $p = .767$, $n = 25$, $r = -.177$; $U(STI) = 48$, $p = .318$, $n = 23$, $r = -.266$).

We calculated the monkeys' peering time and total number of peers into the box (Figures 6 and 7). Monkeys peered significantly longer into the box in both inconsistent conditions (Property/kind and Spatio-temporal) compared to the consistent conditions (Property/kind: Wilcoxon test: $z = 3.462$, $n = 18$, $p < .001$, $r = .816$; Spatio-temporal: $z = 3.632$, $n = 21$, $p < .001$, $r = .793$).

There were no significant differences in the total peering time between the Searching-first and the Looking-first groups across all conditions (Mann-Whitney-U: $U(PKC) = 55.5$, $p = .152$, $n = 19$, $r = .332$;

$U(\text{PKI}) = 57.5, p = .169, n = 26, r = -.272; U(\text{STC}) = 105, p = .134, n = 25, r = .307; U(\text{STI}) = 55, p = .563, n = 23, r = -.129$).

Monkeys peered more often into the box in both inconsistent conditions compared to the consistent conditions (Wilcoxon test: Property/kind: $z = 3.281, n = 18, p < .001, r = .773$; Spatio-temporal: $z = 3.352, n = 20, p < .001, r = .750$).

There were no significant differences between the Searching-first and the Looking-first groups in the number of peering bouts for each condition (Mann-Whitney-U: $U(\text{PKC}) = 61, p = .058, n = 19, r = .482; U(\text{PKI}) = 58, p = .186, n = 26, r = -.277; U(\text{STC}) = 90, p = .501, n = 25, r = .154; U(\text{STI}) = 57, p = 0.648, n = 23, r = -.108$).

Correlation of Different Measures

We used a Spearman's rank test to determine whether looking measures in the looking time paradigm were correlated to reaching and peering measures in the manual search task, using difference scores.

Looking time and reaching time measures were not significantly correlated (Property/kind conditions: $r = .174, p = .553, n = 14$; Spatio-temporal conditions: $r = .056, p = .863, n = 12$). Similarly, there was no significant correlation between looking time and peering time (Property/kind condition: $r = .358, p = .208, n = 14$; Spatio-temporal: $r = .035, p = .914, n = 12$).

Figure 4

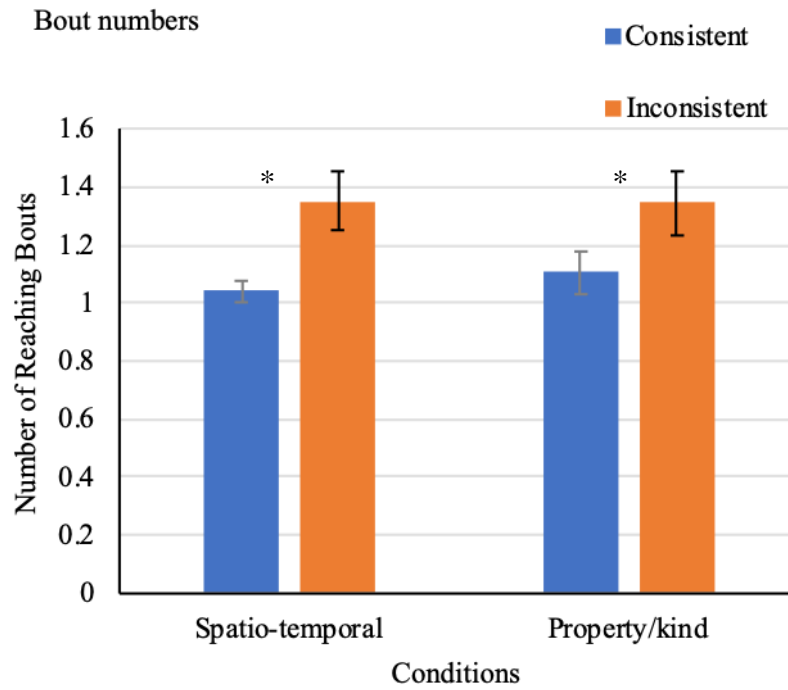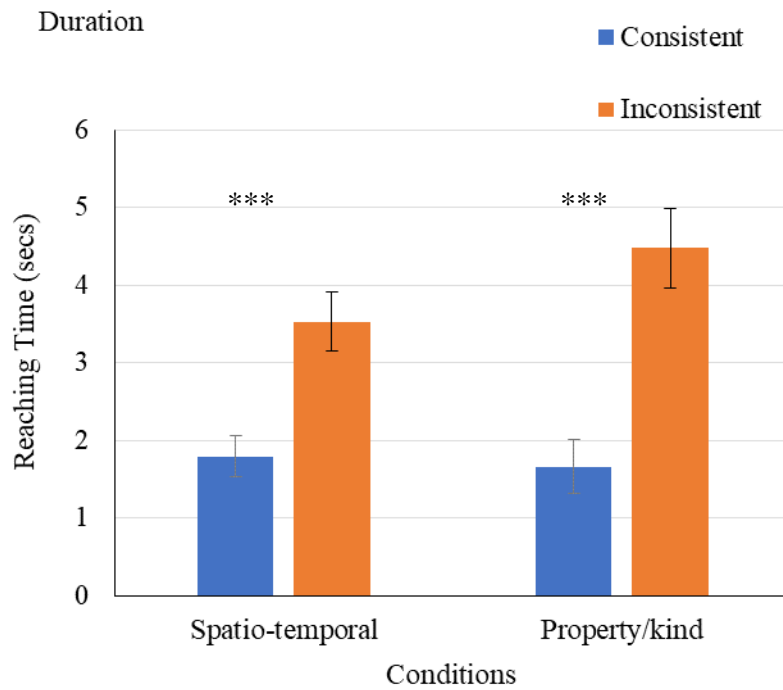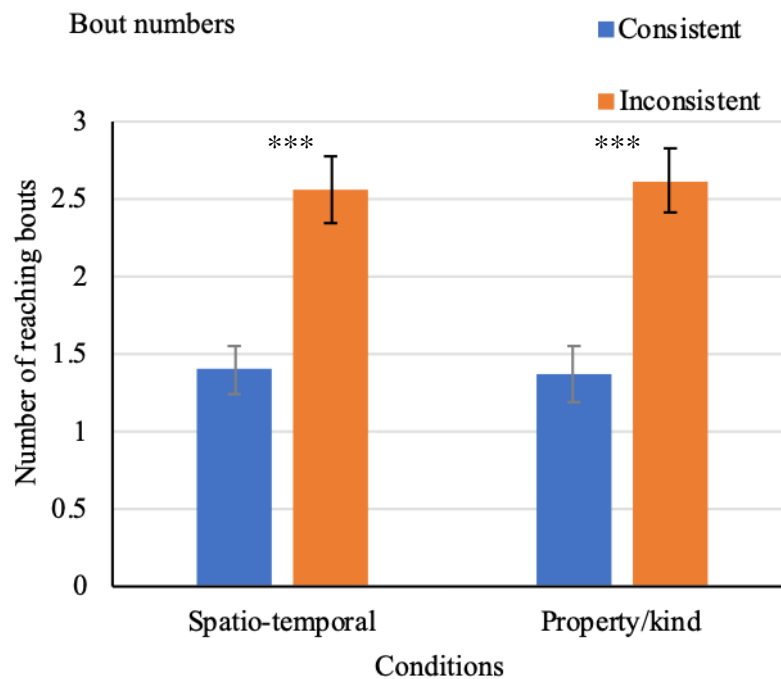*Mean Reaching Time (Secs) Into the Box for Each Condition*
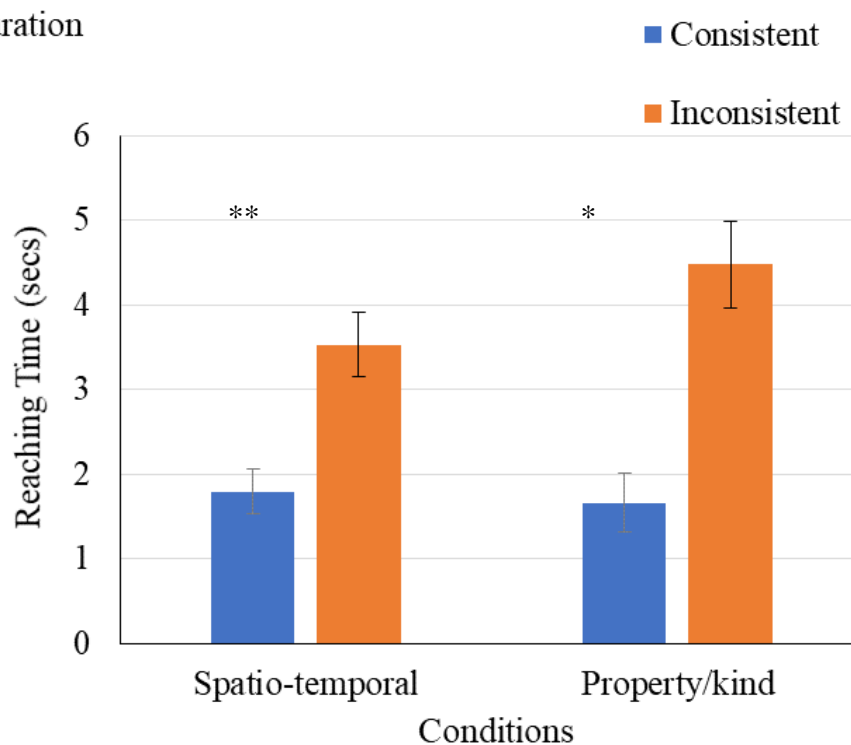
Figure 5

Mean Number of Reaching Bouts for Each Condition

Figure 6

Mean Peering Time (Secs) Into the Box for Each Condition

Figure 7

Mean Number of Peering Bouts for Each Condition

Discussion

Like 12-month-old human infants, great apes and rhesus macaques (Carey & Xu, 2001; Mendes et al., 2008; Santos et al., 2002), capuchin monkeys, a New World monkey species, used both property/kind information and spatio-temporal information to individuate objects. Therefore, our findings align with previous nonhuman animal research to cast further doubt on the notion that using property/kind information to individuate objects depends on language. Furthermore, the finding of co-occurrence of both skills also casts doubt on the idea that they rely on different representational systems. Spatio-temporal and property/kind individuation may share at least some basic cognitive underpinnings (Carey & Xu, 2001; Xu, 2007; Xu & Carey, 1996).

Our study is the first to utilize both looking time and manual search measures in the same subject group with the same experimental apparatus in nonhuman primates, and we found that results from both measures converged at the group level: As a group, monkeys both looked longer and searched more often when the object they found at the bottom of the box was not the same object they saw go in, and when they found only one object after seeing two placed inside. Therefore, our findings suggest capuchin monkeys have a robust competence in using spatio-temporal and property/kind information to individuate objects.

However, at the individual level, the looking and searching performances did not correlate with each other, i.e., the monkeys that looked longer in the visual task did not tend to also reach more in the manual search task. This lack of correlation may not be surprising: if we assume that all of the monkeys can individuate objects using both sources of information, then it follows that the variation in their performance is likely due to noise factors such as motivation or attention. In addition, the limited variability allowed by the reaching measure may have made it difficult to detect any actual correlation in performance. The reaching task typically varied between one or two reaches, whereas the variation in

looking time was greater and more continuous. It may be that, with a more complex task with a richer dependent variable, we could have detected meaningful variation in skill.

Xu and Carey claim that between 10 and 12 months of age, there is a language-driven shift from possessing essentially one sortal concept (bounded physical object) to the possession of sortal concepts for particular kinds of object (such as ducks and balls). However, a similar shift or dissociation has not yet been found in nonhuman animal studies, either in development or in evolution. Various tested species, including apes, macaques, capuchin monkeys, and even domestic dogs and chicks, passed both Spatio-temporal and Property/kind tasks (Bräuer & Call, 2011; Fontanari et al., 2011, 2014; Mendes et al., 2008; Phillips & Santos, 2007; Santos et al., 2002; Uller et al., 1997). This pattern did not change when age was considered in apes (Mendes et al., 2008, 2011) and when young chicks were tested (Fontanari et al., 2011, 2014). Thus, it is less likely that comparative studies did not find the shift because it only appears in ontogeny. Whether this shift exists in phylogeny as well as in ontogeny remains an open question, but the evidence so far suggests that it does not exist.

We have found no dissociation between the ability to individuate objects using property/kind information and the ability to individuate objects by spatio-temporal information in capuchin monkeys. However, this leaves the question of whether there are one or two systems, and when and how they evolved, still open. It is certainly possible that other, as yet untested species show only spatio-temporal, but not property/kind-based individuation and the key point in evolution when the latter ability emerged awaits to be found. It could be argued that capuchin monkeys (like macaques, and great apes) may be a special case amongst primates since they are large-brained extractive foragers that use tools (Ottoni & Izar, 2008; Visalberghi & Tomasello, 1998). Their dietary niche might have provided selective pressure for a different, more advanced, kind of object knowledge that is shared with other extractive foragers or tool-using species, for example nonhuman and human apes, but not necessarily all primates. One interesting direction for future research would be to explore object individuation across a broader number of species from among the primate order, including dietary specialists and generalists, perhaps including different properties available for individuation.

Alternatively, object individuation by property/kind information might have a long evolutionary history, and we would expect it to be widespread. As we discussed, there is evidence to suggest that other animals, evolutionarily even further removed from humans, are also sensitive to property and kind information (Bräuer & Call, 2011; Fontanari et al., 2011, 2014). A final possibility is that both systems of individuation evolved simultaneously and, although temporarily dissociable in ontogeny, they are never dissociated in phylogeny. This suggests that research into an even wider range of species would be illuminating.

A final question is whether our results bear on representations of object 'kinds' at all: it could be argued that capuchin monkeys in the current study showed only the ability to use featural information in object individuation. Further work could examine whether, like apes, capuchin monkeys privilege 'essential' feature of objects (such as what kind of food) over non-essential features (such as color), when individuating objects (Cacchione et al., 2016). However, any featural difference between two objects presented singly and sequentially in a short space of time might cause an observant viewer to anticipate two objects, and so this work might require a different procedure than the one used in the current study, in which some (non-essential) transformation to the object could plausibly have occurred.

The kinds of differences that prompt subjects to infer the presence of two rather than one object seems a critical direction for future research in both human infants and nonhuman animals (Cacchione et al., 2012, 2013). One key function of sortal concepts is to provide persistence conditions across changes in the world. For example, a lump of clay may change shape while retaining its identity, while a squashed statue disappears and only the lump of clay remains. But in the real world, sometimes an object of one sort "transforms" into an object of another sort: for example, a caterpillar becomes a butterfly, or a ball of dough becomes a loaf of bread. Information about transformations of this kind need not be built into the sortal concept. The sortal "caterpillar" is different from the sortal "butterfly" and does not incorporate the information that sometimes caterpillars transform into butterflies. This is an additional bit of background information about the world that children learn separately after they come to possess the sortal concepts

"caterpillar" and "butterfly," and which enables them to expect that in this case, an object of one sort may emerge transformed into an object of a different sort.

This opens up the possibility of a *Sortal Constancy, Knowledge Gain Hypothesis* to explain the different performance of 12-month-olds and 10-month-olds in the Xu and Carey (1996) paradigm. The infants may possess exactly the same sortal concepts, but 10-month-olds may lack the relevant background knowledge that ducks do not change into balls, or vice versa. On this view, the development of object individuation would be much more continuous and need not involve two systems: the difference between 10-month-olds and 12-months-olds is essentially just that 12-month-olds have learned a bit more about the world. Future research might examine in more detail what causes participants to use a contrast between two objects presented singly and sequentially to infer the presence of a single object that has undergone a transformation, rather than two objects.

Conclusion

One essential function of any system of folk physical understanding of the world is object individuation and tracking. Our study indicates that, when faced with the problem of object individuation, capuchin monkeys demonstrate sensitivity to both spatio-temporal and property/kind information, and therefore supports extant evidence that property/kind individuation is not distinctive to language users (humans), and extends the phylogenetic distance within primates in which the ability is found. In addition, we found that capuchins display their sensitivity to spatio-temporal and property/kind information by both looking time and reaching time measures, which suggests that it is a robust ability used both for perception and action. Ours is the first study to test the same individuals in the same object individuation tasks with both measurements. The convergence of results we found at the group level is more compatible with the idea of a single cognitive system for object individuation operating both for perception and action. However, our results are not radically incompatible with the possibility of two systems of representation that is traditionally linked to dissociations in performance. It could be argued that the two putative systems are simply well established in the adult members of the species we studied and therefore they failed to dissociate.

Finally, we argued that the type of sensitivity to property and kind information uncovered by the paradigm we used is compatible both with the presence and with the absence of sortal concepts and essentialist beliefs. We therefore conclude that, despite the considerable empirical evidence accumulated in recent years, Tinklepaugh's (1928) view that there is no "true evidence" of the "representative factors" underlying object individuation remains correct: many theories remain compatible with the evidence, and only further experimental and theoretical work can provide a fuller picture.

Acknowledgement

References

Berthier, N. E., DeBlois, S., Poirier, C. R., Novak, M. A., & Clifton, R. K. (2000). Where's the ball? Two- and three-year-olds reason about unseen events. *Developmental Psychology*, *36*(3), 394–401. https://doi.org/10.1037/0012-1649.36.3.394

Bonatti, L., Frot, E., Zangl, R., & Mehler, J. (2002). The human first hypothesis: Identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology*, *44*(4), 388–426. https://doi.org/10.1006/cogp.2002.0779

Bräuer, J., & Call, J. (2011). The magic cup: Great apes and domestic dogs (*Canis familiaris*) individuate objects according to their properties. *Journal of Comparative Psychology*, *125*(3), 353–361. https://doi.org/10.1037/a0023009

Cacchione, T., & Burkart, J. M. (2012). Dissociation between seeing and acting: Insights from common marmosets (*Callithrix jacchus*). *Behavioural Processes*, *89*(1), 52–60. https://doi.org/10.1016/j.beproc.2011.10.010

Cacchione, T., Hrubesch, C., & Call, J. (2013). Chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*) quantify split solid objects. *Animal Cognition*, *16*(1), 1–10. https://doi.org/10.1007/s10071-012-0545-3

Cacchione, T., Hrubesch, C., Call, J., & Rakoczy, H. (2016). Are apes essentialists? Scope and limits of psychological essentialism in great apes. *Animal Cognition*, *19*(5), 921–937. https://doi.org/10.1007/s10071-016-0991-4

Cacchione, T., & Rakoczy, H. (2017). Comparative metaphysics: Thinking about objects in space and time. In J. Call, G. M. Burghardt, I. M. Pepperberg, C. T. Snowdon, & T. Zentall (Eds.), *APA handbook of comparative psychology: Perception, learning, and cognition* (pp. 579–599). American Psychological Association. https://doi.org/10.1037/0000012-026

Cacchione, T., Schaub, S., & Rakoczy, H. (2012). Fourteen-month-old infants infer the continuous identity of objects on the basis of nonvisible causal properties. *Developmental Psychology*, *49*(7), 1325–1329. https://doi.org/10.1037/a0029746

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object files and object tracking. *Cognition*, *80*(1–2), 179–213. https://doi.org/10.1016/S0010-0277(00)00154-2

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*(5), 1235–1244. https://doi.org/10.1111/j.1467-7687.2011.01074.x

Fontanari, L., Rugani, R., Regolin, L., & Vallortigara, G. (2014). Use of kind information for object individuation in young domestic chicks. *Animal Cognition*, *17*(4), 925–935. https://doi.org/10.1007/s10071-013-0725-9

Fujita, K., Kuroshima, H., & Asai, S. (2003). How do tufted capuchin monkeys (*Cebus apella*) understand causality involved in tool use? *Journal of Experimental Psychology: Animal Behavior Processes*, *29*(3), 233–242. https://doi.org/10.1037/0097-7403.29.3.233

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways tor perception and action. *Trends in Cognitive Sciences*, *15*(1), 20–25. https://doi.org/10.1016/0166-2236(92)90344-8

Hauser, M. D. (2001). Searching for food in the wild: A nonhuman primate's expectations about invisible displacement. *Developmental Science*, *4*(1), 84–93. https://doi.org/10.1111/1467-7687.00152

Hood, B., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events: Two-year-olds fail to reveal knowledge of solidity and support. *Child Development*, *71*(6), 1540–1554. https://doi.org/10.1111/1467-8624.00247

Hood, B., Cole-Davies, V., & Dias, M. (2003). Looking and search measures of object knowledge in preschool children. *Developmental Psychology*, *39*(1), 61–70. https://doi.org/10.1037/0012-1649.39.1.61

James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: An fMRI study. *Brain*, *126*(11), 2463–2475. https://doi.org/10.1093/brain/awg248

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, *41*(3), 841–849. https://doi.org/10.3758/BRM.41.3.841

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing "what" and "where" systems. *Trends in Cognitive Sciences*, *2*(1), 10–18. https://doi.org/10.1016/S1364-6613(97)01113-3

Macdonald, C., & Whiten, A. (2011). The "Living Links to Human Evolution" research centre in Edinburgh Zoo: A new endeavour in collaboration. *International Zoo Yearbook*, *45*(1), 7–17. https://doi.org/10.1111/j.1748-1090.2010.00120.x

Mendes, N., Rakoczy, H., & Call, J. (2008). Ape metaphysics: Object individuation without language. *Cognition*, *106*(2), 730–749. https://doi.org/10.1016/j.cognition.2007.04.007

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*(3), 407–414. https://doi.org/10.1007/s10071-010-0375-0

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia*, *46*, 774–785. https://doi.org/10.1016/j.neuropsychologia.2007.10.005

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*. *5*(7), 309–315. https://doi.org/10.1016/S1364-6613(00)01682-X

Ottoni, E. B., & Izar, P. (2008). Capuchin monkey tool use: Overview and implications. *Evolutionary Anthropology*, *17*(4), 171–178. https://doi.org/10.1002/evan.20185

Penn, D. C., & Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: A comparative, critical review. *Annual Review of Psychology*, *58*(1), 97–118. https://doi.org/10.1146/annurev.psych.58.110405.085555

Phillips, W., & Santos, L. R. (2007). Evidence for kind representations in the absence of language: Experiments with rhesus monkeys (*Macaca mulatta*). *Cognition*, *102*(3), 455–463. https://doi.org/10.1016/j.cognition.2006.01.009

Phillips, W., Shankar, M., & Santos, L. R. (2010). Essentialism in the absence of language? Evidence from rhesus monkeys (*Macaca mulatta*). *Developmental Science*, *13*(4), F1–F7. https://doi.org/10.1111/j.1467-7687.2010.00982.x

Povinelli, D. (2000). *Folk physics for apes: The chimpanzee's theory of how the world works*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198572190.001.0001

Rakoczy, H., & Cacchione, T. (2019). Comparative metaphysics: Evolutionary and ontogenetic roots of essentialist thought about objects. *Wiley Interdisciplinary Reviews: Cognitive Science*, *10*(5), e1497. https://doi.org/10.1002/wcs.1497

Santos, L. R., & Hauser, M. D. (2002). A non-human primate's understanding of solidity: Dissociations between seeing and acting. *Developmental Science*, *5*(2), F1–F7. https://doi.org/10.1111/1467-7687.t01-1-00216

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*(3), 241–264. https://doi.org/10.1016/S0010-0277(02)00006-9

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605–632. https://doi.org/10.1037/0033-295X.99.4.605

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology*, *13*(2), 113–142. https://doi.org/10.1111/j.2044-835X.1995.tb00669.x

Tinklepaugh, O. L. (1928). An experimental study of representative factors in monkeys. *Journal of Comparative Psychology*, *8*(3), 197–236. https://doi.org/10.1037/h0075798

Uller, C., Xu, F., Carey, S., & Hauser, M. (1997). Is language needed for constructing sortal concepts? A study with nonhuman primates. *Proceedings of the 21st Annual Boston University Conference on Language Development*, (1995), Vol. 21 665-677.

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*(3), 249–280. https://doi.org/10.1207/S15327647JCD0103_1

Visalberghi, E., & Tomasello, M. (1998). Primate causal understanding in the physical and psychological domains. *Behavioural Processes*, *42*(2), 189–203. https://doi.org/10.1016/S0376-6357(97)00076-4

Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition*, *72*(2), 125–166. https://doi.org/10.1016/S0010-0277(99)00035-9

Wilcox, T., & Baillargeon, R. (1998). Object individuation in infancy: The use of featural information in reasoning about occlusion events. *Cognitive Psychology*, *37*(2), 97–155. https://doi.org/10.1006/cogp.1998.0690

Wilcox, T., & Chapa, C. (2002). Infants' reasoning about opaque and transparent occluders in an individuation task. *Cognition*, *85*(1), B1-B10. https://doi.org/10.1016/S0010-0277(02)00055-0

Wilcox, T., & Schweinle, A. (2002). Object individuation and event mapping: Developmental changes in infants' use of featural information. *Developmental Science*, *5*(1), 132–150. https://doi.org/10.1111/1467-7687.00217

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, *85*(3), 223–250. https://doi.org/10.1016/S0010-0277(02)00109-9

Xu, F. (2007). Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2007.08.002

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, *30*(2), 111–153. https://doi.org/10.1006/cogp.1996.0005