The representational systems for object and agent in new world monkeys

Da Zhang

A thesis submitted for the degree of PhD
at the
University of St Andrews

2021

Full metadata for this item is available in
St Andrews Research Repository
at:
https://research-repository.st-andrews.ac.uk/

Identifier to use to cite or link to this thesis:

DOI: https://doi.org/10.17630/sta/727

This item is protected by original copyright

Candidate's declaration

I, Da Zhang, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 50,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in August 2016.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 20/05/2021 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date 20-5-21 Signature of supervisor Juan C. Gómez

Amanda Seed

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Da Zhang, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date 20/05/2021 Signature of candidate

Date 20-5-21 Signature of supervisor

Juan C. Gómez

Amanda Seed

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Da Zhang, understand that by declaring that I have original research data or digital outputs, I should make every effort in meeting the University's and research funders' requirements on the deposit and sharing of research data or research digital outputs.

Date 20/05/2021 Signature of candidate

Permission for publication of underpinning research data or digital outputs

We understand that for any original research data or digital outputs which are deposited, we are giving permission for them to be made available for use in accordance with the requirements of the University and research funders, for the time being in force.

We also understand that the title and the description will be published, and that the underpinning research data or digital outputs will be electronically accessible for use in accordance with the license specified at the point of deposit, unless exempt by award of an embargo as requested below.

The following is an agreed request by candidate and supervisor regarding the publication of underpinning research data or digital outputs:

No embargo on underpinning research data or digital outputs.

Date 20/05/2021 Signature of candidate

Date 20/05/2021 Signature of supervisor

Juan C. Gómez

Amanda Seed

Acknowledgements

I would like to thank my supervisor, Dr Amanda Seed and Dr Juan-Carlos Gomez, for the discussions, comments, and advice all through my PhD. Thanks for your patience, support, and encouragement during these five years.

I would also like to thank the members of staffs in the school of Psychology and Neuroscience, and the members of the Scottish Primatology Research Group. Awesome research communities and wonderful people. Special thanks to Dr. Justin Ales, Prof. James Anderson, Prof. Dick Byrne, Prof. Josep Call, Dr. Cat Hobaiter, Dr. Innes Jentzsch, Prof. Dave Perret, and Prof. Andy Whiten for the advice, comments, discussion, or supports. Thanks to the University of St Andrews for the opportunity to study in this prestigious and beautiful university.

Many thanks to my colleagues, the students and research fellows in our laboratory, Christoph, Elisa, Elle, Laras, Lucy, Raphaelle, Zeynep, etc. I am grateful to spend five years with these brilliant people in such a diverse, friendly (and funny), and collaborative atmosphere. I would also like to thank the other researchers and staffs for the great time we spent at the zoo, including Amanda (the keeper), Callum, Donald, Donner, Hazel, Kate, Liz, Matthias, Patience, Roslin, and a lot of other names I couldn't list all here. Thanks for your help and company during the memorable time in the zoo.

Thank the Living Links to Human Evolution Research Centre at the Royal Zoological Society of Scotland, Edinburgh Zoo, for the opportunity for me to conduct monkey research. Thank you for the support. Many thanks to the squirrel monkeys and the capuchin monkeys, those who participated or those who tried to disturb my experiments. You are both adorable fluffy fellows and I will definitely miss you.

Thanks to my family, especially my parents. you are the first mentors that lead me to start the journey of scientific research. Not to mention your love and support. Also, thanks to my partner Xiaoyu, who is also pursuing a PhD. We accompany and support each other.

Last but not least, many thanks to my country and people, not only for funding my current PhD study, but also for twenty years of public education I went through before this PhD.

Funding

This work was supported by the China Scholarship Council (201603780022).

Table of Contents

Abstract

Representing the environment in its most basic components, namely objects and agents, is a fundamental feature of human cognition which we may share to different extents with nonhuman animals. This thesis explored some manifestations of these abilities in two new world monkey species, squirrel monkeys and capuchin monkeys.

We first investigated squirrel monkeys' ability of individuating object by spatiotemporal and property/kind information with a "magic box" paradigm using both manual search and looking time measures (chapter 2). The squirrel monkeys failed both tasks with both measures, whereas capuchin monkeys showed individuating competence with exactly the same tasks and apparatus in a previous study.

Chapter 3 tested and explored the possibility that squirrel monkeys failed the "magic box" tasks that capuchin monkeys passed because they acted so fast that they didn't form or use this type of object representations to guide their actions. In fact, in a touchscreen-based object tracking/catching game (Whack-a-cricket task), the squirrel monkeys were slower to "catch" a moving "cricket" compared to capuchin monkeys.

In chapter 4, we tested squirrel monkeys with another individuation task that included two separate barriers instead of a single box. The squirrel monkeys preferred to search the last-visited-location first when either spatiotemporal or property/kind information suggested that only one object was present. This preference disappeared when either information indicated that there were two objects, one behind each barrier. We conclude that squirrel monkeys are therefore able to individuate objects using both kinds of information when tested with an appropriate task.

In the last chapter, we investigated whether capuchin monkeys can locate a causal agent based on an event that he initiated from a hidden location. Capuchin monkeys located the hidden agent when they saw an object pushed, raked, rolled, or thrown across a table seemingly by the experimenter behind one of two screens (agentive trials), but not when they saw an object roll down a ramp or fall off a block after a shake of the table (arbitrary control trials) that contained no information about the agent' location. This result suggests capuchin monkeys can use motion events to infer the location of a causal agent, an ability also demonstrated by human infants.

Taken together, our studies show that squirrel monkeys and capuchin monkeys have some core abilities to represent some of the fundamental properties of objects and agents, comparable to those demonstrated by human infants, which supports the core knowledge view that such representational systems may have a long evolutionary history and exist widely in primates.

Keyword: Object individuation, squirrel monkey, capuchin monkey, causal agent

Chapter 1 General introduction

Most scholars agree that the basic building blocks of human cognition, representations of objects, space, number, and other agents, are likely to be shared with other primates if not more broadly with other vertebrates (Haun, Jordan, Vallortigara, & Clayton, 2010; Hauser & Spelke, 2004; Santos & Hood, 2012; Seed & Tomasello, 2010; Spelke & Kinzler, 2007; Tomasello & Call, 1997; Vallortigara, 2012). These shared basic cognitive skills and representational capacities help primates navigate the world and cope with similar evolutionary challenges in their environments. But there is also broad agreement that human cognition has some unique features, such as a unique capacity for abstract reasoning or for explicitly representing other's minds (Call & Tomasello, 2008; Seed & Tomasello, 2010; Tomasello & Herrmann, 2010; Weiss & Santos, 2006). Delineating the features of our cognition that are shared, and those that are unique has been a focus for comparative psychologists.

One basic application of these core cognition skill sets is to represent the environment in its most fundamental structure, namely objects and agents (Rakoczy & Cacchione, 2019; Xu & Carey, 1996). With such knowledge and ability, an individual can form internal cognitive symbols to represent external reality, conceive the persistent existence of these entities, keep track of them even when they are temporarily unperceivable, and use these representations for higher level cognitive processing. For example, when I received my kindle a month later after I sent it for repair service, I could recognize this was the kindle I sent, and appreciated the work of the technician who lives in another country. This example illustrates a daily occurrence for human adults. We sometimes lose track of an object, but we know it is somewhere in the environment and we can still recognize it when it reappears. We can work with other people at distance and give credit to them without direct face-to-face interaction. We are able to do all these things owing to our representational systems.

It is an interesting question for psychologists where the roots of the underlying representational abilities are. Developmental researchers have tried to answer this question by proposing a series of so called "core knowledge systems", each serving a special collection of purposes (e.g., Spelke & Kinzler, 2007). At least four core systems have been proposed so far. The first one is an object representational system centred on a group of spatiotemporal principles such as cohesion, solidity, and continuity. Under these principles, an object should be a bounded whole-piece of entity that moves (after contact) coherently along a continuous path. This system serves to represent inanimate objects and their physical movement. The second system represent agents and their actions. In addition to following the same spatiotemporal principles as objects (e.g., boundedness and continuity), an agent is able to propel itself or other entities intentionally towards a goal, usually by an efficient means. The third system represents numbers in an imprecise, abstract way, whilst it enables comparison and combination of numbers. The fourth system represents the geometry of the environment, including the distance, angle, and relations among the extended surface layouts. These core systems emerge early in human ontogeny, and probably serving as building blocks of human cognition. Nonhuman animal studies also have been inspiring and enriching this theoretical framework by shedding light on the evolutionary roots of these core cognition modules (Hauser & Spelke, 2004; Vallortigara, 2012; Weiss & Santos, 2006).

For instance, in the previous kindle example, at least two systems are required, the one for representing objects (the kindle), and the one for representing agents or social partners (the technician). These systems have been proposed to be largely innate, at least in their most basic architecture (Baillargeon & Carey, 2012; Santos & Hood, 2012; Spelke, Breinlinger, Macomber, & Jacobson, 1992; Spelke & Kinzler, 2007). Both systems can be found in very young human infants. An infant can understand a covered toy is still there at the age of 4 months, even if they cannot manually retrieve it until around 9 months (Spelke, Kestenbaum, Simons, & Wein, 1995), realize an object moving out from a barrier is the same one which just went behind that barrier (e.g., Xu & Carey, 1996), or infer a hidden experimenter to be the cause of an event (e.g., Saxe, Tenenbaum, & Carey, 2005; Saxe, Tzelnic, & Carey, 2007).

Comparative psychology, on the other hand, focuses on the phylogenetic origin of these representational systems. Non-human animal studies, especially nonhuman primate studies, provide evidence that these core systems may also widely exist in the animal kingdom. For example, apes, old world monkeys such as rhesus macaques, and new world monkeys such as capuchin monkeys, demonstrate object permanence, object individuation , and some understanding of other’s mental states, such as intentions and goals (Call & Tomasello, 2008;

Jaakkola, 2014; Krupenye & Call, 2019; Rakoczy & Cacchione, 2019). These findings suggest that the core systems have a long history in the evolutionary history. But how far can we trace back the origin of these systems? The ancestor of primates? Are the core knowledge systems in nonhuman animals the same as the ones we humans have, namely, serving for the same specific purposes and showing the same signature limits? Both questions need extended research into more non-human primate species. The comparative research looking into such questions would not only enrich our understanding about the natural history of cognition, but also help characterize the innate "core" of the systems and shed light on the mechanism of these representational systems.

The current thesis aims to focus on the object and agent representation skills of two cebid monkey species, the capuchin and squirrel monkey, especially the latter one. These two species are both new world monkeys whose ancestors diverged from ours about 35 million years ago (e.g., Gómez, 2004). Thus, in phylogeny they are more distant relatives of humans compared to great apes (diverged about 3 million years ago) and old-world monkeys (diverged about 20 million years ago). Whereas numerous experimental studies have revealed object and agent representational competence in apes and old-world monkeys such as rhesus macaques, it would be natural to extend these experiments to new world monkeys. Among them, capuchin and squirrel monkeys are an interesting pair of species to be studied and compared. On the one hand, these two "sister" species share a lot of common ecology. They both live in the rainforest of central/south America. Their habitat largely overlaps, and the two species have been reported to form mixed species troops in the natural environment (e.g., Leonardi, Buchanan-Smith, Dufour, MacDonald, & Whiten, 2010). And they both live in cohesive groups. Both species have omnivorous diets, preferably fruits and insects, and large brain-body-ratio (Fragaszy, 1985; Fragaszy, Visalberghi, & Fedigan, 2004; Roth & Dicke, 2005). On the other hand, there seem to be some potential differences in terms of their cognitive skills despite of their phylogenetic relatedness. One of the most dramatic differences is that tool use and flexible object manipulation have been observed in the wild capuchins but not squirrel monkeys. For an example related to object representation, capuchins are able to pass the invisible displacement tasks of object permanence (Mathieu, Bouchard, Granger, & Herscovitch, 1976; Schino, Spinozzi, & Berlinguer, 1990), whereas the squirrel monkeys can only pass visible displacement tasks but cannot understand the invisible displacement, which is considered to be cognitively more complex and appears later in the ontogeny of apes and humans (de Blois, Novak, & Bond, 1998). Although squirrel monkeys are a species that can be widely found in zoos and laboratories as well as capuchin monkeys, the research on squirrel monkey's cognition is relatively sparse compared to the capuchins. The latter have shown competence in many other cognitive areas, especially related to agents. For instance, capuchins are sensitive to another individual's intentions and goals (Drayton & Santos, 2014; Phillips, Barnes, Mahajan, Yamaguchi, & Santos, 2009). No similar finding has been reported in squirrel monkeys insofar. Therefore, capuchin monkey and squirrel monkey seem to become a very interesting pair of subject species, considering their place in the phylogenetic tree, the similarity in their ecology, the reported differences in cognition, and the wide distribution of these two species in research sites.

In the rest of this chapter, we will introduce two lines of experimental studies, one in each cognitive domain. On the object representation domain, we will focus on object individuation, the ability to identify and track objects in the environment. Whereas on the agent representation domain, we will explore the ability to infer the presence of a hidden agent based on the events caused by its action. Existing studies testing these two abilities in human infants and nonhuman primates will be reviewed to depict a broad picture and thereby to suggest how the four experiments in the current thesis contribute to the field.

1.1 Representing objects: object individuation

Object individuation refers to the ability to parse visual input arrays into discrete individual objects, determine the number of objects present in the scene and keep track of them across space and time (Baillargeon et al., 2012; Fontanari, Rugani, Regolin, & Vallortigara, 2014; Santos & Hauser, 2002; Xu & Baker, 2005). This ability enables an individual to address questions such as "what (objects) are there?" or "is it the same object(s)?". For instance, a foraging monkey needs to parse fruit from the tree, track the food while reaching for it, and sometimes re-identify the food after it was out of sight because of occlusion for a moment. And if, for example, the monkey accidentally drops a fig that falls through the branches to the ground, the monkey has to track its movement to be able to search for it in the fallen leaves. Would the monkey know exactly what she is searching for and how many targets there are (how many figs fell to the ground)? More specifically, in the case above does the monkey expect to find a single fig, or would she be satisfied if she found a grape or half a fig? Is her search guided by a representation of what she is looking for?

Previous research suggests that at least two types of potential information sources may be helpful to individuate objects. First of all, we individuate objects based on spatiotemporal information. A pile of stuff moving as a bounded whole with a clear and connected path can be viewed as an "object" (Spelke et al., 1995; Spelke & Kinzler, 2007). Second, one may also be able to rely on property or kind information, using some specific features (e.g., shape, colour, size, etc.) to distinguish between different objects and track their individual identity, or to represent objects by their kind categories on a more conceptual level (Mendes, Rakoczy, & Call, 2008; Rhemtulla & Xu, 2007; Wilcox, 1999; Wilcox & Baillargeon, 1998a; Wilcox & Chapa, 2004; Xu, 2002; Xu, Carey, & Quint, 2004; Xu, Carey, & Welch, 1999). Testing spatiotemporal and property/kind object individuation respectively and examining if they occur separately or concurrently in human development and primate phylogeny could enlighten us on the mechanisms of the object representation system.

1.1.1 Object individuation in human infants

In humans, the ability to individuate objects by either spatiotemporal or property/kind information seems to emerge early in infancy according to developmental studies during the past thirty years. In a seminal study, Xu and Carey (1996) presented infants with a set of occlusion events in which objects may move out or behind the screen(s). In the spatiotemporal task, two screens were displayed with a gap between them, so that the motion path between these two screens was visible. Infants would see a toy (i.e., duck) emerge from one side, turn back and emerge from the other side several times. The motion could either be continuous or discontinuous depending on whether the object appeared in the gap of two screens. After the demonstration, the screens were lowered down to reveal either one or two ducks. The result showed that in the discontinuous condition, 10 months old infants looked longer if there was only one duck, while they looked longer when there were two in the continuous condition, suggesting the infants individuated the objects based on the spatiotemporal principles. However, infants of the same age failed to show any surprise in a similar set up which included one screen and two different toys (i.e., a duck and a ball). After watching a ball and a duck emerge from the different sides of the screen alternatively, 10-

month-olds looked longer when 2 objects were revealed, suggesting that they were expecting a single object to be revealed when the screen was lowered. These results suggested that not until 12-months can infants individuate object by their featural differences, indicating a developmental gap between using spatiotemporal versus property/kind information.

Since Xu and Carey's research (1996), convergent evidence has been provided for the developmental shift between 10- and 12-months (Krøjgaard, 2000; Van De Walle, Carey, & Prevor, 2000; Xu, 1999; Xu et al., 2004, 1999). Van de Walle and colleagues replicated the study by using a manual search measure (Van De Walle et al., 2000). After a similar display as Xu and Carey's, the infant could search into a box. Twelve but not ten months old infants searched longer when they only found one toy (unexpected outcome) in the property/kind trials, indicating the violation of their expectation based on property/kind representations. In another experiment using a different task, the infants had to segregate an ambiguous array consisting of a pair of adjacent objects (for example, a duck mounted on a car) when only the upper one or both objects were lifted (Needham & Baillargeon, 2000; Wilcox & Baillargeon, 1998a; Xu et al., 1999). Again, 12-month-olds showed surprise when both objects were lifted while 10-month-olds would only be surprised after watching the objects moving separately before the test, i.e., with spatiotemporal cues (Needham & Baillargeon, 2000; Xu, 1999; Xu et al., 1999).

What does this developmental gap between spatiotemporal and featural individuation mean? Xu and Carey argued that these results reflect a developmental change in the ontogeny of language and sortal concepts. Sortal refers to concepts of kinds, which provide principles of individuation (Xu, 2007). For example, the suitable question for an adult in Xu and Carey's experiment would be "how many DUCKs are there behind the screens" or "is the BALL the same as the DUCK?" These all-capital words here stand for sortal nouns mapping basic-level sortals that enable the process of individuation. In contrast, these questions may be ambiguous if the basic-level sortal nouns are replaced by more general sortals (e.g., OBJECT). On the other hand, one may be able to use spatiotemporal information with the general sortal concept "OBJECT". This can still yield the correct answer to "how many OBJECTs" in a spatiotemporal task, although the difference between objects of different kinds is not specified by the concept OBJECT. Hence Xu and colleagues (Xu, 2007; Xu & Carey, 1996; Xu et al., 2004) suggest the OBJECT-first hypothesis —younger infants understand only a global concept of bounded physical object defined with a set of spatiotemporal rules. Evidence using looking time measures reveal that infants as young as 4 months may already do this (Spelke et al., 1995). But it is suggested that not until 12 months can infants represent the basic-level sortal concepts which are critical for the success in the property/kind tasks.

Considering that infants start language acquisition by the end of the first year, Xu and Carey suggested that the conceptual development of sortal is associated with language learning. The first supporting evidence was the positive correlation between noun comprehension and performance in the property/kind task (Xu & Carey, 1996). The number of nouns the infants comprehended predicts success in the property kind tasks. Moreover, learning sortal nouns may be necessary in individuating objects. Evidence suggests that preverbal infants benefit from labelling (Xu, 2007; Xu, Cote, & Baker, 2005). For example, 9-month-olds succeed as well as 12-month-olds in the classic property/kind task if the experimenter says “look, a duck” and “look, a ball” when the object emerges from the screen. This effect still occurred with unfamiliar objects and nonsense words, but not with distinct expressions, tones, or sounds (Rivera & Zawaydeh, 2007; Xu, 2002).

More interestingly, infants search longer in a box when the content is labelled with two distinct names compared to a single label twice, suggesting that distinct labelling could guide the representation of objects as belonging to different kinds (Xu et al., 2005). These findings suggest that infants may interpret words as sortal nouns, which map onto the representations of different kinds. Therefore, the words might serve as “essence placeholders” which allow the infants to form more elaborated representations and corresponding expectations of one or two objects in the tasks (Xu, 2002; Xu et al., 2005). Or, as suggested in some stronger claims, language somehow enables the initial representations sortals (the exact role of language is still unclear. For some discussion on the possible explanations, see Xu, 2002; Xu, 2007).

1.1.2 Two systems for object representation?

To summarize, Carey and Xu proposed a two-system account (hereafter Carey-Xu account) to explain the developmental shift between 10- and 12-month-old infants (Carey & Xu, 2001). At the age of 10-month, infants only have the first object representations system that is based on spatiotemporal principles, whereas at the age of about 12-month infants develop the second, property or kind-based representational system, which relies on language acquisition and sortal concepts.

Carey and Xu used an integrated framework combining object indexing and object-file theories to explain the first system (Carey & Xu, 2001; Xu, 1999). When representing an object, an index attends to the object and tracks its spatiotemporal “address”, just as a finger points at a target object. This so-called FINST (FINgers of INSTantiation, Pylyshyn, 2001) serves as the initial phase of an object-file, the representational symbol of the object. Thus, the object representation is organized based on spatiotemporal principles since the beginning. When a new object appears from nowhere or a violation of these principles happens (for example, an object appears successively on two locations between which no spatiotemporal continuous path can be found), the individual establishes a new object-file and therefore represents a new object. The first system then indexes and tracks this object across space and time by storing and updating the spatiotemporal information in the object-file. In this process, the primacy of spatiotemporal information is one of the key signature limits of this system. Namely, a new object file can only be established based on spatiotemporal information, but not property/kind change. Moreover, though properties may be bound to an existing object-file, they only play a secondary role in the service of individuation. Hence, ten-month old infants who could individuate objects based on spatiotemporal information failed to individuate objects by property/kind (Xu & Carey, 1996). The so-called ‘tunnel effect’ is another example, when an object A moves into a barrier and a different object B shortly reappears from the other side of the barrier, even human adult subjects tend to represent only a single object, regardless of the property/kind difference (Burke, 1952; Carey & Xu, 2001; Flombaum, Kundey, Santos, & Scholl, 2004; Scholl & Pylyshyn, 1999).

According to this framework, we can find the commonality between this first system and the core object representational system. Both theories describe a system that emerges early in ontogeny to represent bounded physical objects based on spatiotemporal information whereas conflicting property/kind information will be overridden (Carey & Xu, 2001; Spelke & Kinzler, 2007). Therefore, the first system underlies the spatiotemporal individuation in 10-month-olds but not the property/kind individuation in 12-month-olds. The latter may rely on a second, kind-based system (Carey & Xu, 2001). According to the “Linguistic Sortal

Concept Hypothesis" that Xu and colleagues defend, this second system establishes object tokens on a conceptual level based on their kind with the help of sortal nouns serving as labels, whereas the first system represent object with mid-level object-files (Carey & Xu, 2001).

However, this second system remains relatively unclear. Carey and Xu didn't specify the capacity and constraints of this second system with as much detail as the first system in their theoretical papers (Carey & Xu, 2001). They use an example to illustrate the existence and mechanism of the second system. When watching panel 1 and 2 in Figure 1-1 successively, one may experience apparent motions that the black rabbit on the top left seemingly "turns" into the white bird on the top right and the inverse change happens on the bottom. The first, spatiotemporal system takes charge in this situation, minimizing the total amount of movement considering the spatiotemporal arrays but not the property/kind information. On the contrary, if one takes a long-term interval (say, 5 mins) between watching these two panels, it is more plausible to imagine diagonal movements of a black rabbit and a white bird. In this context, the second system identifies the objects based on their kind membership. This example seems to shed some light on the mechanism and scope of application of the two object representation systems. Yet it is difficult to apply the same rationale to interpret the empirical evidence from which the Carey-Xu account derived. In Xu and Carey's property/kind object individuation tasks, and also in the follow-up studies we mentioned above, 12-month-olds did individuate the two objects by property/kind information, despite that fact that the timing is more similar to the case in which the first system is in charge in the example above. Moreover, according to Carey and Xu's own theory, property information can be bound to the object-file generated by the first system, while it only plays a secondary role in individuation and can be overlooked under high task demands. Thus, it is possible that with a longer break between watching panel 1 and panel 2, the property/kind information bound to the object-file is more likely to "survive" and to be taken into account, rather than being ignored under the time pressure. In this interpretation, only the first system would be sufficient.

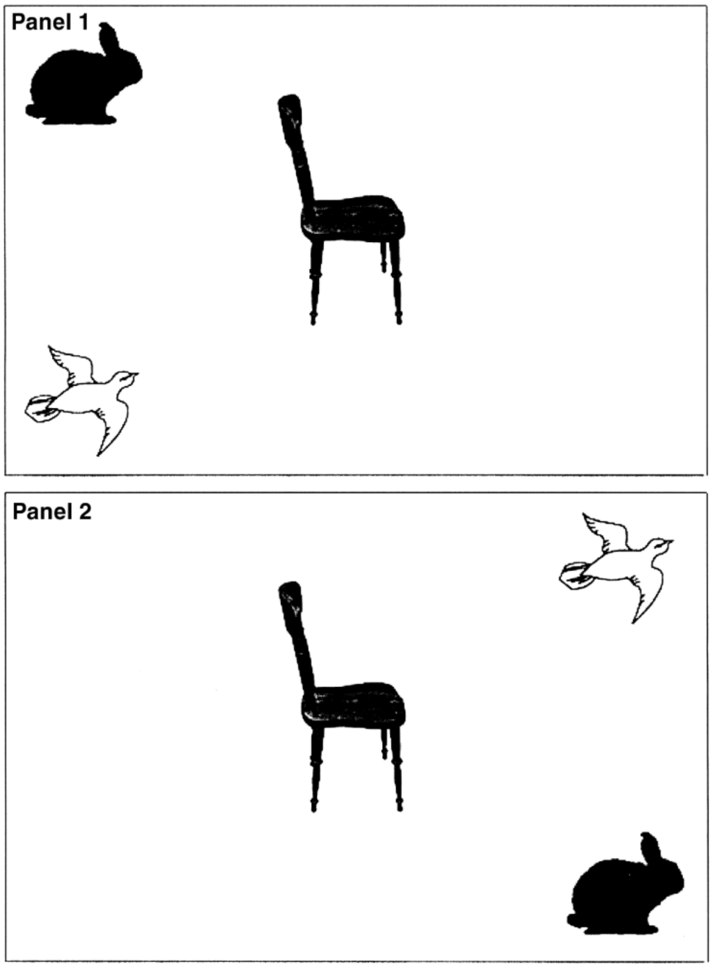


Figure 1-1 A “prima facie” example to illustrate the different mechanism of the two object individuation systems used by Carey and Xu (2001)

Another version of the two-system account integrates the object indexing system with the “where” and “what” visual systems (Leslie, Xu, Tremoulet, & Scholl, 1998; Xu, 1999). The two visual systems have been differentiated and discussed from a physiological perspective since half a century ago (Schneider, 1969). Both originated from the primary visual cortex, the “where” system refers to a dorsal pathway involving the parietal lobe, processing the location information, whereas the “what” system that is responsible for properties, such as colour and shapes, refers to a ventral circuit including the temporal lobe that is responsible for properties, such as colour and shapes (Goodale & Milner, 1992; Milner & Goodale, 2008; Schneider, 1969). Critically in Xu and Leslie’s theory, the object indexing system can receive the input from the “where” system since an early age, which accounts for 10-month-olds successful spatiotemporal object individuation, while not until the end of the first year can an individual integrate the two visual systems, so that one can individuate objects by property/kind (Leslie et al., 1998). The developmental change between 10- and 12-month-olds reflects a maturation process, in which language may play merely a scaffolding role of "labelling" different categories with specific sortal nouns, or an integration process in which language plays a more pivotal role of bridging separate representational systems. The latter role of language is inspired by a hypothetical idea in Spelke's core knowledge systems theory (Spelke & Kinzler, 2007).

We consider this latter version of the two-system account more like a one-system-plus account, though it underscores a second representational system based on property/kind information ("what" system). The Spelke's core object representation system based on spatiotemporal principles is the central mechanism of object individuation, namely THE one system. This system answers the questions of "which or WHAT?" and "where" with the FINST and object-file, whereas the "what" system plays a secondary role to encode property/kind information which awaits to be bound to the object-file (Carey & Xu, 2001; Leslie et al., 1998; Xu, 1999). In other words, the success of property/kind individuation relies on the integrative effort of both systems, rather than solely the output of the "what" system. In contrast, in the original version, i.e., the Carey-Xu account we mentioned earlier, the kind-based system seems to work independently to the spatiotemporal system (Carey & Xu, 2001).

In summary, researchers posit there is more than one object representational system to explain the empirical evidence of a developmental gap that 10-month-olds individuate objects only by spatiotemporal information whereas 12-month-olds can further individuate objects by property/kind. A critical interpretation is that this gap reveals a stage-like change, which lead to several versions of two-system account (Carey & Xu, 2001; Xu & Carey, 1996). Regardless of the differences across these versions, the advocates of the two-system account acknowledge two hypotheses in common. First, the OBJECT-first-hypothesis highlights an early developed or even innate object representational system centred on spatiotemporal principles and representing only "Spelke" object(s) (Spelke, 1990; Spelke et al., 1995; Xu, 1999). Given the signature limit known as spatiotemporal primacy, representing property/kind information requires another representational system. Before the second, kind-based system matures (or/and becomes ready to work together with the first system), one would not be able to individuate objects by property/kind. Second, the language gain hypothesis assumes that language is important to the second system (Carey & Xu, 2001; Xu, 2002; Xu & Carey, 1996; Xu et al., 2005). The acquisition of language, especially sortal nouns, may help an individual to label objects and therefore categorize objects by kinds. A radical form of this hypothesis argues that the linguistic sortal concept plays not only a scaffolding role, but also a pivotal role in the kind-based object representational system. In addition to these two hypotheses, there is a question remaining for both versions of the two-system account. How do we individuate objects by property/kind? Or, what exactly is the second system that the Carey-Xu account posits? It is not only that researchers find difficulties to depict the second system, but there is also evidence that questions the stage-like gap between spatiotemporal and property/kind individuation, the very base of the two-system accounts.

1.1.3 Evidence of early individuation success in preverbal human infants

There are two lines of research that challenge the idea of a stage-like gap between individuating by spatiotemporal information and by property/kind individuation. The first one came from developmental studies that revealed early competence of property/kind individuation. Wilcox and colleagues (e.g., Wilcox, Alexander, Wheeler, & Norvell, 2012; Wilcox & Baillargeon, 1998a, 1998b; Wilcox & Chapa, 2004; Wilcox & Schweinle, 2002) used simplified tasks to demonstrate that preverbal infants can individuate objects using property/kind information only (for reviews, Baillargeon et al., 2012; Needham & Baillargeon, 2000; Stavans, Lin, Wu, & Baillargeon, 2019).

Wilcox, Baillargeon, and collaborators first questioned the failure in 12-month-olds by suggesting a difference between two task types: “event mapping” versus “event monitoring” (e.g., Wilcox & Baillargeon, 1998a, 1998b; Wilcox & Schweinle, 2002). The former one, taking Xu and Carey’s standard task as an example, requires a subject to represent the target object(s) in the occlusion event during habituation, but mapping the knowledge of object representation onto a non-occlusion event during the test phase when the occluder is lowered. Wilcox and Baillargeon (1998a, 1998b) suggested that it was the “mapping” across events that introduced extra cognitive demands for the infants and led to failure of individuation. In contrast, they provided two lines of evidence that in a single ongoing event, infants would be able to individuate object by property/kind. For instance, when object A moved into a screen and a different object B subsequently moved out from the other side, 4.5-month-olds would look longer if the screen was too narrow to conceal both objects at the same time, compared to when the screen was wide enough or when the two objects were the same (Wilcox, 1999; Wilcox & Baillargeon, 1998a, 1998b; Wilcox & Schweinle, 2002). The so-called "narrow-screen task" saves the step of the non-occlusion event, and arguably implies that younger infants can individuate objects based on property/kind information when they only need to monitor the event (for an argument against this interpretation, see Carey & Xu, 2001). A more interesting finding was reported in a task using a transparent screen versus no screen, apart from the opaque screen used in the standard tasks (Wilcox & Chapa, 2002). The two objects emerge successively from the opposite side of an opaque screen, and then this opaque screen is lowered to reveal the outcome with either a transparent screen in front of it or no screen. Intriguingly, 9.5-month-olds were surprised at the inconsistent outcome (two different objects appeared but only one left) only in the transparent screen condition but not when there was no screen. Baillargeon and colleagues interpret this finding as a support to their suggestion that infants can individuate object by property/kind in an ongoing event: when the opaque screen is lowered to reveal a transparent screen, it is still an occlusion event (Stavans et al., 2019; Wilcox & Chapa, 2002). Thus, removing the extra cognitive demands for understanding and remembering the events across different set-ups, infants were able to demonstrate their competence at an earlier age.

In addition, Wilcox and colleagues further showed that by shortening the event sequence younger infants could also succeed in property/kind tasks (Baillargeon et al., 2012; Krøjgaard, 2003; Wilcox & Baillargeon, 1998a; Wilcox & Schweinle, 2002). Instead of objects moving into and out from the occluding screens reversely several times, these shortened tasks included less reversal or even only a one-way movement from one side of the stage to the other. In the latter case, 5.5-month-olds were surprised to find no object behind the lowered barrier after watching an egg move in and a column move out (Wilcox & Schweinle, 2002). The preverbal infants in this study seemed to take the egg and the column as two individual objects, though the spatiotemporal information suggested a single object. Carey and Xu (2001) argued that this success was based on insufficient spatiotemporal information conveyed in the shortened sequence. We found this argument very unconvincing. According to the Carey-Xu account they suggest in the very same paper, the property/kind difference is inadequate for an individual equipped with only the first system to open a new object-file. Thus, the individual who saw an egg went in that barrier would have (only) an object-file of the egg but shouldn't represent the column as a second object when it moved out. In other words, the lack of spatiotemporal information cannot account for the successful use of property/kind information. Meanwhile, Wilcox and Schweinle also used a "single trajectory" version of the spatiotemporal task to test 5.5-month-olds (Wilcox & Schweinle, 2002). The positive results suggest the spatiotemporal information seems to be sufficient for object individuation at this age. Hence, this single trajectory object individuation experiment suggested not only that human infants are able to individuate objects by property/kind earlier than Xu and Carey claim, but also that they encode the property/kind information in such a robust way that it survives the conflict with the spatiotemporal information.

Intriguingly, the "single trajectory" version of the task can be further related to another group of studies that lead to quite opposite empirical findings and interpretations. While the single trajectory tasks we introduced above reduce the cognitive demands to allow an individual to reveal the ability of property/kind individuation, the "tunnel effect" reveals a difficulty to override the spatiotemporal bias and to use property/kind information (Burke, 1952; Flombaum et al., 2004; Flombaum & Scholl, 2006). It is easy to find the similarity between the single trajectory task and the task of the "tunnel effect": an object A moved into an occlusion and an object B came out shortly afterwards. However, in the latter case the subjects (human infants, adults, and rhesus monkeys) overlooked the property/kind difference between the object that went into the barrier and the one that came out, and encoded the event as a single object moving continuously in time and space behind the barrier (Burke, 1952; Carey & Xu, 2001; Flombaum et al., 2004; Flombaum & Scholl, 2004; Scholl & Pylyshyn, 1999; Wilcox & Chapa, 2004). In other words, in very similar (if not exactly the same) experimental contexts in which the continuous path (spatiotemporal information) suggests a single object, but the property/kind difference suggests two different objects, young infants sometimes encode the property/kind information (single trajectory experiments) whereas human adults sometimes do not (Burke, 1952; Flombaum & Scholl, 2006). We agree with Carey and Xu (2001) that the tunnel effect reveals the primacy of spatiotemporal information (over property/kind information) in object representation (also see, Flombaum, Scholl, & Santos, 2012). However, it is noteworthy that the primacy of spatiotemporal information is only one of the signature limits of the first system, and it should not be used as evidence of the existence of a second object representation system based on property/kind or sortal concepts. The fact that human adults sometimes fail to use property/kind information in object representation casts further doubts on this hypothetical second system.

Moreover, while some advocates of the Carey-Xu account acknowledge that preverbal infants can individuate objects by featural or property information, they still argue that kind-based individuation is a different ability that emerged much later (e.g., Carey & Xu, 2001; Xu, 2007; Xu, Carey, & Quint, 2004; Xu, Carey, & Welch, 1999). However, there has also been evidence that preverbal infants can individuate objects by its kind or category, with or even without preverbal priming (Stavans & Baillargeon, 2016; Wilcox, Woods, & Chapa, 2008). For an example of priming, Stavans and Baillargeon (2016) presented the 4-month-old infants with different pairs of object-action (e.g., a pair of tongs was used to pick things up, whereas a masher compressed things) before testing them with a standard object individuation task. The infants' success suggests that the "functional demonstration" beforehand helped the infants to encode the objects as tools of distinct categories and therefore improved their performance (Stavans & Baillargeon, 2016). For another example that without priming, Kibbe and Leslie (2019) found that infants remembered the categorical identity but not the perceptual feature of a hidden object. This and other evidence emphasized the difference between featural and categorical information in object individuation, and further suggest that the categorical (rather than featural) information can be used in object representation at an age earlier than Xu and Carey predicted (Bonatti, Frot, & Mehler, 2005; Bonatti, Frot, Zangl, & Mehler, 2002; Kibbe & Leslie, 2019; Surian & Caldi, 2010).

Given the mixed results provided by researchers from different sub-fields using various experimental paradigms, it is still an unresolved issue the question of how objects are individuated based on spatiotemporal and property/kind information. Notably, after decades of research, there are still new theoretical accounts arising. In a recent paper, Stavans, Baillargeon, and collaborators have proposed a new model to explain these mixed results (Stavans et al., 2019). This new model also suggests two early-emerging systems for object individuation, but with quite different meanings: (1) an object-file (OF) system that makes use of spatiotemporal and categorical information; and (2) a physical reasoning (PR) system that takes causal relevant features into account. We find at least two major differences between this "OF+PR" model and the Carey-Xu account. First of all, while an object-file system is included in both models, its function is different. The new model assumes that the

OF system itself can encode an object's kind/category, so that the "kind-based system" Xu and Carey suggested is no longer necessary. The empirical findings showing that preverbal infants succeed in different-categories individuation tasks seem to support this assumption (Bonatti, Frot, Zangl, & Mehler, 2002; Kibbe & Leslie, 2019; Surian & Caldi, 2010). Second, the new model assumes that the OF and PR systems work together in parallel to represent object(s) in physical events, and individuation failures are due to disagreement between systems, rather than due to the late development of one of the systems, as Xu and Carey's stage-like two-system account suggested (Stavans et al., 2019).

Stavans et al. (2019) further make two more assumptions (or rather two "patches," as the authors used them to explain two exceptions) in an attempt to explain the existing conflicting results across the different infant tasks with their new model. First, the object representation of an individual or his/her prediction of an event depends on the timing of when the individual needs to respond. During an event, the individual relies on the output of the PR system rather than that of the OF system, whereas when the event comes to an end, the answer depends on whether and how the two systems agree with each other. Second, the nature of how the PR and OF systems disagree after the event ends matters: if the two systems disagree in what the authors called a "qualitative" way (0 vs. 1, or 'nothing' vs 'something'), the infant can resolve the disagreement and represent one object, whereas if the disagreement is "quantitative" (e.g., 1 vs. 2), younger infants will fail to form the correct representation. The first assumption can be used to explain the infants' early success in event-mapping tasks, taking these tasks as ongoing occlusion events. The second assumption can be used to explain the findings in those single-trajectory tasks, in which there will be either one or no object behind the screen (or in the container). Note that in Xu and Carey's standard task the two different objects moved back to the same screen so that there would be either one (according to the OF system based on spatiotemporal information) or two (according to the PR system based on featural information) objects, namely a quantitative disagreement that young infants fail to resolve.

This novel OF+PR account is aimed to be a comprehensive theory within the field of developmental psychology that tries to explain both the results that support Xu and Carey's two-system account and the results against it. However, it does not address another line of study stemming from comparative psychology, that has yielded evidence that may be incompatible with the Carey-Xu account.

1.1.4 Evidence from nonhuman animal studies

A basic presumption of comparative psychology is that an ability found in human adults that emerges in early life is likely to evolve in our ancestors (Hauser & Spelke, 2004; Weiss & Santos, 2006). Therefore, one can expect to find such ability in some of our close relative species, such as nonhuman primates. In terms of object individuation and object representation, returning to the example scenario in the beginning that a monkey looked for a dropped grape, the monkey certainly does not understand or use language: it cannot label the previous piece of food as “grape” and the currently found piece as “apple” to differentiate them. Whereas infant researchers proposed a stage-like Carey-Xu account based on the findings of a developmental gap between ten- to twelve-month-old, it is a question for comparative researchers how these non-linguistic individuals will perform on object individuation tasks, especially on the property/kind individuation tasks. This question is especially crucial for Xu and Carey’s account which underscores the role of language (Xu, 2007; Xu & Carey, 1996).

Uller and colleagues (1997) replicated Xu and Carey’s original study in rhesus monkeys. The monkeys looked longer when during the familiarization phase they saw a squash slice and a carrot slice emerge alternately from behind the screen, but only one piece of food was revealed in the test phase, indicating that the monkeys were sensitive to property/kind information (Uller, Xu, Carey, & Hauser, 1997). Munakata and colleagues replicated another experiment of Xu (1999) on rhesus monkeys (Munakata, Santos, Spelke, Hauser, & O’Reilly, 2001). The results showed that the monkeys expected two adjacent but distinct objects to move separately, therefore indicating an individuation ability based on featural information. Santos and colleagues (2002) found similar results with another task based on Van de Walle et al.’s study, which used a box and containment event (“box paradigm/task” hereafter) instead of barrier(s) and occlusion events (Van de Walle et al., 2000). Two types of box tasks were designed (Santos, Sulkowski, Spaepen, & Hauser, 2002). One was a box version of the original task of Xu and Carey’s (or a manual search version of Uller’s task), in which the food item(s) was taken out and shown to the monkey before being inserted back to the box.

For instance, in a different-object property/kind task, the experimenter took out an object A and put it back, followed by a similar sequence but with an object B. The monkeys searched longer around the box when the demonstration involved two different objects but only an object A was found (violation condition), compared to the consistent condition, where the same kind of object was taken out and inserted back. Another type of task included only the inserting, in analogy to the single-trajectory task in infant studies. This type of task has been dubbed the modified "remainder task" by Stavans, Baillargeon and colleagues (Stavans et al., 2019). Namely, after an object A was inserted into the box and the subject retrieved a different object B in the property/kind-violation condition, the question addressed by the subject could be glossed as "whether there is any object remaining in the box?". In this example, a monkey with property/kind individuation ability should expect an object A to be left in the box and search for it. Thus, for both types of tasks, only one piece of food could be found in a trial. The researchers calculated searching time as the total time spent looking and reaching into the box after the only piece of food was retrieved. The monkeys searched longer in the violation trials when the outcome of searching was unexpected in terms of quantity (spatiotemporal condition) or quality (property/kind condition). Furthermore, Phillip and Santos (2007) tested the same species controlling for the perceptual features of the object stimuli. A piece of identical "food" (actually a white plastic piece) was removed from either an apple or a half coconut: it looked as if the experimenter cut off a piece from the fruit, and then inserted it into a box. The plastic chunk was surreptitiously removed before the monkey could search in the box which was preloaded with a real piece of fruit. The monkeys searched longer if they found a piece of food of the unexpected kind, showing kind understanding about differences (Phillips & Santos, 2007).

Mendes and colleagues (2008) tested three great ape species with a modified version of Santos and colleagues' task (Phillips & Santos, 2007; Santos et al., 2002), using a stricter measure of reaching time. On the one hand, Mendes et al (2008) considered only the behaviours of searching into, but not around the container as in Santos and colleagues' studies. This measure avoids the danger of confounding the time spent searching for the expected but missing object with the looking time in response to violations (Mendes et al., 2008; Santos et al., 2002). On the other hand, Mendes et al included the subjects' searching behaviour AFTER they found the expected food, i.e., the "second search". It was suggested that if the searching was driven by a clear expectation of the missing food, the subject would continue searching when the first piece of food was retrieved but should stop searching (or search less often) after the expected food was given (Mendes et al., 2008). The results showed that apes were more likely to keep searching in the box after they retrieved an unexpected kind of food. Mendes et al argued this new measure could rule out an alternative explanation that the animals search longer in total simply because of an increased arousal level provoked by more food items (spatiotemporal conditions) or diverse kinds of food items (property/kind conditions) that might lead to more searching effort. The stringent measure used here indicated that apes formed an expectation of a certain object based on property/kind individuation.

Kersken and colleagues tested capuchin monkeys with a box task similar to that used in Mendes et al.'s experiment (Kersken, Zhang, Gomez, Seed, & Ball, 2020). Both manual search and looking time measures were used by manipulating the bottom panel(s) of the box, which was either a panel with an access hole in the manual search trials, or a transparent glass in the looking time trials. The capuchin monkeys looked longer or searched more often when two objects were dropped in and only one was found (spatiotemporal violation) or when a grape was dropped in, and a date was found (property/kind violation). These convergent findings across measures suggest that capuchin monkey can individuate objects by both types of information.

The ability to individuate objects based on property/kind information has not only been found in primates. Bräuer and Call tested apes and dogs with a "magic cup game" in which the food placed in a cup might be surreptitiously substituted. Both apes and dogs showed "surprise" (more begging and looking for the apes, and more smelling for the dogs) in a comparable way in the substituted trials (Bräuer & Call, 2011). There is also evidence that birds might succeed in property/kind tasks (Fontanari et al., 2011, 2014). In one experiment (Experiment 2 of Fontanari et al., 2014), the researchers presented newly hatched chicks with 2 identical screens behind which either social stimuli (imprinting objects) or food items were hidden. On one side, an imprinting stimulus and a mealworm came out and hid behind a screen consecutively, while on the other side a single worm or imprinting stimulus appeared and returned behind the screen twice. Given that chicks tend to approach the larger group of items, if they had property-kind knowledge, they should show a preference for the side with two different objects. The results supported this prediction. Moreover, the chicks didn't show a similar preference when choosing between two identical objects and two different objects, both pairs presented simultaneously (Experiment 3), ruling out the alternative explanation that the chicks simply preferred a more varied group of stimuli. It is noteworthy that this study measured preferential choice (choosing between two alternatives) instead of cumulative behaviours, such as looking time, searching time, or the number of visits. That is, subjects reacted before any outcome had been revealed, whereas in most of the other nonhuman animal studies, reactions (looking or reaching) to an expected or unexpected outcome were measured. Such preferential choice measure may have an advantage according to the OF+PR model, since the subjects made the decision DURING the event so that it is more likely they correctly represent the objects based on their PR system (Stavans et al., 2019).

In summary, comparative researchers have studied several species including rhesus monkeys, apes, dogs, and domestic chicks (Bräuer & Call, 2011; Kersken et al., 2020; Fontanari et al., 2011, 2014; Mendes et al., 2008; Santos et al., 2002; Uller et al., 1997). All these nonhuman animal species individuated objects by property/kind information without language. Such findings challenged a strong version of Xu and Carey's language gain hypothesis (Carey & Xu, 2001; Xu, 2002, 2007). Moreover, the subjects could use both spatiotemporal and property/kind information to individuate objects. Thus, no dissociation between these two kinds of individuation was found, as predicted by the Carey-Xu account. Compared to the mixed results in human infants, these nonhuman animal studies provide evidence inconsistent with Xu and Carey's account from a comparative-evolutionary perspective. Yet, one could argue that the evidence is still scarce; only a few species have been tested and it is still possible that at least one species can be found whose performance corresponds to the dissociation between systems suggested by Carey and Xu (Carey & Xu, 2001). This possibility encourages comparative researchers to expand the object individuation study to more species.

In relation to the recent model of Stavans and collaborators, it is noteworthy that they didn't make any specific predictions about whether or not nonhuman animals can individuate objects (Stavans et al., 2019). Their OF+PR model focuses on the whether or not and how conflict between systems can be resolved to enable successful individuation in human infant development. From a Carey-Xu account perspective, this means that both of the two systems required for object individuation have matured (Carey & Xu, 2001). Given the OF+PR model was published after most of the empirical data of this thesis was collected, it is beyond our scope to test this model in our experiments. Nonetheless, we will consider any implications of our results for this model in the discussion sections.

1.2 Representing agents: who dunnit?

Whereas researchers have been interested in how individuals represent objects outside their immediate perception, as discussed above, a natural parallelism appears to be how individuals also represent "missing" agent(s). "Agent", as well as "object", serves as a fundamental building block of our external world and our internal representations of it. Our social interaction would be almost impossible without the ability for an individual to represent other individuals. From an evolutionary perspective, agent representation is no doubt essential for many species, especially for the social ones as most primates including humans are. It enables an individual to trace a prey to be hunted, to spot an enemy to be avoided, or to recognize a social partner to be collaborated with. These tasks often need to be accomplished when the target agent is temporarily "missing" in one's perceptual field. Thus, the ability to represent an agent can evolve from the adaptive challenges of ones' social life and ecological niche, and be part of the core knowledge systems of cognition. In other words, we expect to find an agent representation ability in nonhuman primates, as the literature reveals for human infants. Before we explore such ability, a question needs to be answered. What is the difference between an agent and an object? Or, perhaps more accurately, what makes an object also an agent? We will begin with a brief discussion of the key feature that characterizes an agent.

1.2.1 Differentiating agents from objects

The concept of "agent" usually refers to animate, goal-directed beings. To some degree, agents can be taken as a special form of object. A typical agent usually conforms to all the key features of a "Spelke Object". Or in other words, an agent, in most cases, follows the spatiotemporal principles that define a "Spelke object": a bounded entity that moves coherently and continuously through space and time (e.g., Aguiar & Baillargeon, 1999; Hespos & Baillargeon, 2001). In some cases, it is tolerable to find an agent that breaks one of these rules. For example, infants would not be surprised to find an agent (such as a human avatar or a self-propelled object) walked into location A and then walked out from location B without the continuous path between A and B, but the agent should not just disappear (Kuhlmeier, Bloom, & Wynn, 2004; Luo, Kaufman, & Baillargeon, 2009). From this perspective, representing an agent shares a similar aspect with representing an object.

Apart from this commonality, from a common-sense perspective, one would argue that an agent has its own characteristics that differentiate it from a simple object. Some experimental evidence has supported this common-sense perspective. For an example from the object individuation literature, human infants differentiate agents from inert objects at a basic level. In an object individuation experiment using Xu and Carey's original version of the task, 10-month-olds successfully identified and counted a humanlike object and a non-humanlike object as two "objects", and they also did so when it was a humanlike object compared to a doglike object (Bonatti et al., 2002). Note that in Xu and Carey's original studies infants at this age could not use property/kind differences to individuate objects (Xu & Carey, 1996). Bonatti et al suggest that infants represent not only the Spelke physical object, as Xu and Carey suggested in their Object-first hypothesis (Xu & Carey, 1996), but also encode some basic-level categorical information to represent humans vs objects (Bonatti et al., 2002). In these cases, the feature used to identify an agent may be the face. Facial features capture an individuals' attention since a very young age (e.g., M. H. Johnson, Dziurawiec, Ellis, & Morton, 1991; S. C. Johnson, Slaughter, & Carey, 1998) and can be informative for agency attribution in human infants (Bonatti, Frot, & Mehler, 2005; Bonatti et al., 2002; S. C. Johnson, 2003) and infant macaques (e.g., Tsutsumi, Ushitani, Tomonaga, & Fujita, 2012).

Nonetheless, static features such as faces may not be necessary for an agent to be recognized. In another object individuation study, 10-month-old infants differentiated an agent (a toy animal, a caterpillar or a bee) from an inert object (a cup or a cube) based on some "dynamic" features: the agent moved autonomously in a non-rigid way, whereas the object was moved in a rigid way by the hand of the experimenter (Surian & Caldi, 2010). In this case, there is no facial feature on the agents for infants to encode. Instead, Surian and Caldi (2010) suggested the infants used the "dynamic features" in the object individuation process. More specifically, it is the self-propulsion feature of agents that matters.

Researchers, since Premack, have suggested that self-propulsion is one of the most prominent features of an agent (Premack, 1990; Premack & Premack, 1997). Agents are capable of self-propelled motion, while objects are not. Thus, subjects tend to take a self-propelled stimulus as an agent and hold different expectation about how it will move compared to an inert object (Baillargeon, Wu, Yuan, Li, & Luo, 2009; Crichton & Lange-Küttner, 1999; Luo & Baillargeon, 2005; Poulin-Dubois, Lepage, & Ferland, 1996). For example, infants by the age of five to six months old are not surprised to see a self-propelled object spontaneously reverse its direction. In contrast, they are surprised to watch these events happen on an inert object, unless there was a reason for it, such as the object bouncing back after hitting a wall (Luo et al., 2009).

Interestingly, some researchers argue that mere self-propulsion may not be sufficient to conclude agency. First of all, infants attribute goals to a typical agent, such as humans. In a Woodward-paradigm, for example, infants who had been familiarized with an actor moving towards one of two target destinations were surprised to find that the actor changed to the other target, suggesting that the infants recognized the goal of this actor (Woodward, 1998). However, researchers using the same paradigm on self-propelled versus inert objects found mixed results. Twelve-month-old infants did not attribute goals to a self-propelled object (S. C. Johnson, Shimizu, & Ok, 2007; Shimizu & Johnson, 2004). But 5-month-olds succeed at the task if they watched the self-propelled object changing trajectory back and forth beforehand (Luo & Baillargeon, 2005). Similar mixed results were found using a slightly different paradigm, in which the self-propelled object moved around an obstacle to approach the target in the familiarization phase. When the obstacle was removed in the test phase, infants looked longer at a straight path than a detour path, if they saw only one fixed path in the familiarization phase, suggesting they didn't take the object as an agent (Kamewari, Kato, Kanda, Ishiguro, & Hiraki, 2005). In contrast, if infants had watched the object make different paths around the obstacle, infants expected the object to move straight to its target as if it were a goal-directed agent (Csibra, 2008). These findings indicate that besides the self-generated motion, the variability of behaviour, or a certain level of unpredictability, may also have an impact on the attribution of agency.

Some recent findings support such explanation. Agents may also move at their will in a more unpredictable manner, while objects do not. For example, when 7-month-olds watched a moving stimulus through a semi-translucent screen (so that the identity of the stimulus remained ambiguous), they would look longer if it was an agent (such as a dog) that moved straight, or if it was an inanimate object (such as a truck) that changes speed and direction while moving (Träuble, Pauen, & Poulin-Dubois, 2014). Similar findings of using speed/direction change to determine agency have also been reported in nonhuman animals (Abdai, Ferdinandy, Terencio, Pogány, & Miklósi, 2017; Di Giorgio, Lunghi, Vallortigara, & Simion, 2021; Rosa-Salva, Grassi, Lorenzi, Regolin, & Vallortigara, 2016).

In sum, human beings distinguish an agent from other physical objects from infancy. Apart from biological features such as faces, a key property of an agent is its potential to change the motion states of its own or other objects. As Leslie indicated in his theory of agents, an agent has "an internal and renewable source of energy or force" that enables it to "act in pursuit of goals or re-act to the environment as a result of perceiving" (Leslie, 1996). Such internal energy allows an agent to be the cause of events. Therefore, one of the central issues in representing an agent is recognize its causal capability and the causal role it may play.

1.2.2 The causal role of agents

Agents are endowed with internal energy, which enables them to be the cause of actions or events, whereas a mere physical object can only be the patient of actions. At least three kinds of events and corresponding causal roles an agent may play have been studied. The empirical findings suggest human infants in their first year are sensitive to these differences and probably recognize these causal roles an agent can play.

First of all, the agent is able to affect itself, changing its motion status without external reasons that objects may need. For example, infants are surprised or even scared to find an object, but not an agent, starts moving on its own (e.g., Poulin-Dubois et al., 1996). Apart from initiating its own motion, agents may also resist external forces, reverse its trajectory, and even float in the mid-air, whereas infants look longer if it is an object doing all these (Luo et al., 2009). As we introduce in the last section, changing the motion speed and direction is also a typical agentive behaviour, a criteria of agency attribution used by not only human infants but also other animals including chicks and dogs (Abdai et al., 2017; Di

Giorgio et al., 2021; Rosa-Salva et al., 2016; Träuble et al., 2014). In this kind of context, agents play both the role of causal reason and recipient.

Secondly, a slightly different scenario in which more researchers have been interested is an agent's goal-directed action *towards* an object, in which the agent, or more specifically, the agent's mental states, serve as the cause of the action. As we mentioned above, infants can attribute a goal to an agent but not an object (Gergely & Csibra, 2003; Kamewari et al., 2005; Luo & Baillargeon, 2005; Luo et al., 2009; Shimizu & Johnson, 2004). In these and other experiments, mostly using the Woodward-paradigm (Woodward, 1998), infants around 6-month-old or even as young as 3-month-old form an expectation about which of two objects an agent prefers and will approach or grasp based on the prior experience during the familiarization phase. Infants were surprised and looked longer if the agent chose another target in the test phase. While one explanation of these results suggests that the infants can represent the causality in such events (e.g., Leslie, 1996), Csibra and collaborators argued that the infants succeed on these tasks based on a teleological but not causal interpretation (Csibra & Gergely, 1998; Gergely & Csibra, 2003). That is, the future state or the outcome of the action is cited to justify the behaviour and account for the event, i.e., the behaviour of approaching is taken as a rational way to achieve the goal of grasping the target object. Whereas the causal interpretation would need to refer to the prior condition that necessitates the behaviour by providing its generative source, in this context, the mental states of the agent such as its desire for the target object (Gergely & Csibra, 1998; Gergely & Csibra, 2003). According to this interpretation, infants may not necessarily represent the causal property of the agent in goal-directed actions. A stronger and more direct evidence of causal agent representation may come from a perspective which concerns the source rather than the goal of an event.

Finally, apart from propelling itself or launching goal-directed actions, an agent can also act *upon* an object, serving as a cause to change that object's states. One of the typical scenarios is the so-called Michottian launching event: a situational agent transfers its internal energy, force, or momentum to the situational patient through physical contact (Michotte, 1947). Early research revealed that infants over half a year old are sensitive to causally relevant spatiotemporal parameters. For instance, they discriminated between events in which the patient object started moving immediately after a collision and those in which the patient object started moving with a delay (Leslie, 1982, 1984). The infants were sensitive to spatial gaps as well as temporal gaps (Muentener, Bonawitz, Horowitz, & Schulz, 2012). Moreover, infants assigned different causal roles to the two objects in a Michottian launching event showing longer looking when a familiarized causal event was reversed (which also implies a change in the role of agent and patient), compared to when it was a non-causal event (with a delay after collision before launching) that was reversed (Leslie & Keeble, 1987). These findings suggest that young infants are able to encode the cause in such events.

Furthermore, more recent studies explore the sensitivity to the likelihood that an agent is the causal source of an event involving object movement. When watching an ambiguous event in which the source of movement was out of sight, young infants would anticipate that an agent was the cause. For example, Saxe and colleagues designed a scenario in which a beanbag flew out from one of two screens onto a stage while the reason was unperceived (Saxe et al., 2005, 2007). Infants as young as 10-month-old would be surprised and look longer when they found an inanimate object (such as a toy truck) behind the screen from which the beanbag flew out. Similarly they looked longer when an agent (a human hand or a puppet) was found behind the opposite screen to the source of movement. They would not be surprised, however, if the agent was revealed from the screen where the object emerged. In other words, after watching a spontaneous motion event of an inert object, infants expected the cause to be an agent. Moreover, infants would not expect an agentive source to account for the motion of a self-propelled puppet instead of a beanbag. According to these findings, infants expect that (1) a spontaneous motion event of an inert object (but not a self-propelled one) requires a cause; (2) the cause can be inferred based on the event even with no direct perceptual evidence; (3) the primary cause of such event should be an agent rather than an inert object.

Besides, infants can also encode the causality in some other events, where an agent also plays the role of cause. For example, Muentener and Carey (2010) used physical state change events (such as collapsing) when testing 8-month-olds. The infants expected a box to collapse or change its colour whilst playing music after contacted by a hand but not by a toy truck (Muentener & Carey, 2010). For another example by Muentener and collaborators (2014), two-year-old toddlers watched a light flash either spontaneously (no explanation for the flashing of the light) or after the experimenter touched it (an observed cause for the flash). In the former case, the toddlers were more likely to predictively look towards the light, when they then watched the experimenter press a button (which had not been previously related to that light). This result suggests that the toddler inferred the causal relation between the spontaneous event (light flashing) and a previously unobserved action (pressing the button) by the agent, the latter of which served as the cause of the event (Muentener & Schulz, 2014). Toddlers expect agents to act deterministically rather than probabilistically (Wu, Muentener, & Schulz, 2016). In Wu and collaborator's experiments (2016), toddlers were first presented with two relationships (cause A led to effect A and cause B lead to effect B) 3 times, and then watched either the same relationships in a switched order (B→B, A→A) or two changed relationships (A→B, B→A). In the latter case, the probability that cause A led to effect A as well as that cause B led to effect B would be 75%, namely the probabilistic condition, rather than 100% in the former case of the deterministic condition. Afterwards, the toddlers would look longer when an agent (hand or puppet) appeared from the apparatus in the deterministic condition, than in the probabilistic condition. In comparison, the toddlers looked equally long in the two conditions when a non-agentive object appeared. These findings suggest that toddlers inferred an unobserved agent, but not a non-agentive object, to account for the deterministic event. Taking these findings together with Saxe and colleague's studies (Saxe et al., 2005, 2007), when infants observe an event occurs without a reason, they tend to and are able to infer an unobserved agent to be the cause. In other words, infants understand the causal role an agent may play, that is, to act upon an inert object and cause the effect they observed. As we argued before, the finding that individuals encode the agentive causal source in an event, especially when the agent is unperceived at some point during the event, provide a stronger form of evidence of agent representation in infancy.

To summarize this section, we discussed the causal role an agent may play across three widely studied contexts: an agent changes its own states, approaches its goal, or acts upon an object. In each of the contexts, the agent always serves as the cause of the event, whereas non-agentive or inert objects may only play the receptive role. The infant studies we reviewed in this section reveal an early ability to represent agents as entities with causal properties, the internal energy that makes things happen.

1.2.3 Representing a causal agent: a comparative perspective

As discussed since the beginning of this chapter, representing an agent is an important and helpful ability for human beings, and probably other animals. In the previous two sections, we briefly reviewed the studies showing that young human infants distinguish agents from mere physical objects, and that young infants assign causal roles to agents in different events. We concluded that (1) one of the key features of agents is its causal properties, or more specifically, its internal energy and the ability to project that energy to initiate events; and (2) that agents can play causal roles in events in which they move themselves, approach and contact with object(s), or/and impact the object by changing its status. Thus, we propose that one direct route to study the ability to represent at agents is to examine whether an individual is able to infer the existence of an agent, while it remains unperceived, based on the perception of its causal effect.

On the other hand, these and other findings suggest that the ability of representing agents emerges early in human development, which is in line with the Core Knowledge Systems Theory (Spelke & Kinzler, 2007). As we mentioned before, the Core Knowledge Systems theory posits the existence of a module or system to represent agents alongside the module or system to represent objects. Moreover, it also posits that the origins of both modules can be traced phylogenetically as well as ontogenetically. That is, the ability to represent agents can be found in our primate relatives as well as in young human infants.

While the landmark ability of representing hidden causal agents has been extensively studied in the developmental literature, our understanding of how nonhuman primates represent causal agents is relatively sparse. To our knowledge, no research has directly addressed this topic with nonhuman primates. Some studies even seem to imply that nonhuman primates cannot do so. Vonk and Subiaul designed a series of tasks to test whether chimpanzees can use the visibility of limbs to predict a human's ability to perform certain physical tasks, in which either the human agent's arm or leg was required to give the chimpanzee food (Vonk & Subiaul, 2009). The chimpanzee would face a choice between two human agents, one with her arms occluded and constrained (Arms Not Visible, ANV) and the other with legs occluded and constrained (Legs Not Visible, LNV). Therefore, if the food tray was on the floor, a chimpanzee should beg the ANV agent when the food trays were on the agents' laps or beg the LNV agent for the food on the floor. Chimpanzees didn’t alter their choice across conditions depending on the visibility of limbs and the requirements of the task (floor-leg or lap-arm), suggesting that the chimpanzees didn’t spontaneously use the visibility of limbs to infer the agent’s capability to launch certain physical events. Vonk and Subiaul further argued that this finding might indicate a global inability in nonhuman primates to make inferences based on unobservable features. This so-called Unobservability Hypothesis proposed by Povinelli and colleagues posits that nonhuman animals have difficulties to reason about abstract entities that cannot be directly perceived, because they do not seek causal explanations (Vonk & Povinelli, 2012; Vonk & Subiaul, 2009). If Vonk and Subiaul’s experiment revealed the true incompetence of apes and Povinelli’s hypothesis is correct, nonhuman primates would be predicted to fail to locate a hidden causal agent, as they are supposed to not only fail to encode the capability of an agent, but also to lack the ability to seek causal explanations of events.

However, a series of studies testing not only human infants but also apes, macaques, and capuchin monkeys supports the idea that nonhuman primates as well as human infants can infer the capability of an agent to initiate events and distinguish between actions performed by “unwilling” versus “unable” humans (Behne, Carpenter, Call, & Tomasello, 2005; Call, Hare, Carpenter, & Tomasello, 2004; Canteloup & Meunier, 2017; Phillips et al., 2009). For example, when an experimenter tried to give the subject a grape, but the grape fell on a ramp and rolled away, or the grape could not go through the hole due to human inability, the subject was more likely to stay in the testing area and interact with the apparatus and the experimenter. In comparison, a subject would tend to stop interacting or even leave the area if the experimenter teased the subject by pulling the grape back when the subject reached out. Besides the sensitivity to the humans’ intention that the subjects show in these studies, they also seem to demonstrate awareness of the capability and constraints of agents. Moreover, accumulating evidence suggests that nonhuman animals can ascribe mental states to others (for reviews, Call & Tomasello, 2008; Krupenye & Call, 2019). Taking a recent false-belief (FB) study as an example, apes were presented with a video including (a) a human agent chases an ape avatar to a hiding location A, (b) the human leaves the scenario, (c) the ape moves from location A to B, (d) the human returns with a stick. The eye-tracking data revealed that the subjects would look more in anticipation to location A, where the human agent falsely believed the ape avatar to be hiding, but not location B in accordance with the subjects' own believes and the reality (e.g, Krupenye, Kano, Hirata, Call, & Tomasello, 2016). The results suggest the apes ascribe the (false) belief of the human agent, presumably as a cause of the agent's behaviour. Interestingly, the apes didn't make such anticipatory looking tendency towards location A in an inanimate control conditions, in which the human avatar was replaced by shapes (Krupenye, Kano, Hirata, Call, & Tomasello, 2017). These findings imply that the apes encode agents in a different way to inanimate objects, and they reason about the mental states of an agent but not an inanimate object, to predict his behaviour. Note that the agent was away and therefore unperceived when he temporarily left the scenario, and therefore one further interpretation of these findings is that the apes can represent the agent and his belief during his perceptual absence. If so, it is possible that nonhuman primates may also be able to represent agents as possessing the physical capability of causing events.

In terms of seeking causal explanations, three lines of studies suggest that nonhuman animals can infer causes based on their effects. First of all, as we just mentioned, theory of mind studies imply that at least apes can infer the mental states of an agent, which serves as the cause of his/her future behaviour. Second, regarding physical causality, apes can infer the location of an object based on its effect (Call, 2004, 2007; Civelek, Call, & Seed, 2020; Hanus & Call, 2008; Völter & Call, 2014; Völter, Sentís, & Call, 2016). For example, an object (e.g., a food item) can support a board and make it tilt. Thus, a chimpanzee chooses an inclined board over a flat one to find the hidden food item (Call, 2007). Similarly, when presented with a balance-beam with a cup on each side, the chimpanzees would prefer the lowered side, reasoning that the food item and its weight made the beam tilt (Hanus & Call, 2008). In these examples, chimpanzees inferred the reason of a physical effect, and represented its hidden cause, i.e., the unobserved object with its relevant causal properties. These and other findings provide evidence that nonhuman primates can traceback an effect to its cause. What we do not know is if nonhuman primates can represent a hidden agent with its causal capabilities to cause an event and this will be one of the aims of the current thesis.

Outside primates, there is evidence that some birds may be able to infer a hidden causal agent. Taylor and colleagues explored this ability in New Caledonian crows (Taylor, Miller, & Gray, 2012). In their study, there was a food container which the birds would approach to retrieve food, while a stick that might poke out from a hiding location was potentially a threat to the foraging birds in this area. In the hidden causal agent (HCA) condition, a human walked into the hiding location before the stick started to probe, after which the human left the hiding location in full sight of the crow. In the unknown causal agent (UCA) condition, the stick poked as in the HCA, but the crow didn't see an agent walk into or out of the hiding. In the subsequent food extraction phase, the crows were more cautious: inspecting more before retrieving food (or even giving up) in the UCA condition. This finding suggests that the New Caledonian crow inferred a hidden causal agent to account for the poking in the UCA condition. If we interpret this finding as an evidence that corvids are able to represent a hidden causal agent, it is possible that we may find similar abilities in nonhuman primates.

Taken together, the existing comparative research suggests that (1) nonhuman animals have some understanding of agents and what agents are capable of, and that (2) nonhuman animals can infer causes based on their effects. We therefore propose that the ability to infer and represent a hidden causal agent, as a landmark agent representation ability, might be found in our close relatives but further empirical studies are needed to test this and improve our understanding of how our representational system evolved.

1.3 The present study

The current study aims to investigate the abilities to represent objects and agents in new world monkeys. We will explore whether monkeys can represent an entity, either objects or agents, along with it's the features that may help to identify that entity. That is, for objects, the property/kind information, and for agents, the causal property to cause an event. Both abilities are fundamental components of the Core Knowledge Systems that need to be traced in evolution. The literature we reviewed in this chapter revealed such early competences in human infants, both in terms of object and agent representation. If we take a core knowledge perspective, one might well expect to find that these abilities originated in evolution and should be present in nonhuman primates.

In the domain of object knowledge, we aim to test whether squirrel monkeys are able to individuate objects based on property/kind information as well as spatiotemporal information. We predict that squirrel monkeys will succeed in the spatiotemporal individuation task, given that every other tested primate species has such ability. In relation to the ability to pass a featural/property individuation task, we do not hold a strong prediction of the results, given there is at least two conflicting theories with different predictions. The Carey-Xu account would predict that it should be possible to find species that can successfully use spatiotemporal but not property/kind information to individuate objects. If we found such pattern in squirrel monkeys, the result would support the Carey-Xu account, and contribute to determine the evolutionary time point when the second system that encodes property/kind information of objects evolved. The more recent Stavans and colleagues' OF+PR account (an object-file system and a physical reasoning system) makes no predictions for comparative research. Our empirical studies of objects individuation will be based on the use of a "magic box" task that we previously used to reveal capuchin monkeys' competence on both types of individuation tasks (Kersken et al., 2020).

In the domain of agent knowledge, we will test whether capuchin monkeys are able to infer the location of a hidden agent as the causal source of observed events. We adapt Saxe and colleague's (Saxe et al., 2005, 2007) beanbag scenario previously used to test infants. That is, between two potential hiding places, a target object moves in a variety of manners, either with or without an observed cause. If an individual is able to represent an agent with its causal capability to initiate the movement of an inert object, this individual shall be able to infer the location of the hidden agent as the cause of the observed event. We do not have specific predictions about the performance of our capuchin and squirrel monkeys in this task, given this is the first experiment to explore this ability in a nonhuman primate species, to our knowledge. However, we tentatively assume that the capuchin monkeys are likely to succeed in this task, given the existing evidence that (1) capuchin monkeys are sensitive to intentional action by agent (e.g., Phillips et al., 2009); (2) capuchin monkeys can make psychological inferences based on human agent' actions and effects (Drayton & Santos, 2014; Tao, 2016); (3) capuchin monkeys can be sensitive to causal information (Fujita, Kuroshima, & Asai, 2003; Visalberghi & Tomasello, 1998; Völter & Call, 2017); (4) not only young human infants but also crows are reported to have succeeded in similar tasks (Saxe et al., 2005, 2007; Taylor et al., 2012).

Chapter 2 Squirrel monkeys' object individuation: Magic Box task

2.1 Introduction

2.1.1 Background

Xu and Carey (1996) proposed a stage-like change underlying the performance difference on spatiotemporal and property/kind object individuation tasks between 10- and 12-month-old human infants, which served as the key evidence for the Carey-Xu account they developed later (Carey & Xu, 2001). Under a combined framework of object-file (e.g., Leslie et al., 1998; Scholl & Leslie, 1999) and object indexing (e.g., Pylyshyn, 2001), this account suggests that, first, an encapsulated perceptual system indexes and tracks bounded physical objects (Spelke object) based on spatiotemporal rules. With this first system, 10-month-olds would possess only a single sortal, roughly glossed as "bounded physical object", which is not sufficient to individuate the two different objects in the property/kind task. Later, a kind-based system develops to individuate objects based on sortal concepts more specific than the "Spelke object", such as "ball" and "duck" (Xu, 2002, 2007). Moreover, Xu and Carey suggested that language acquisition, namely the production of sortal nouns, is the driving force behind the emergence of the kind/property-based system that allows infants to individuate objects as sortal kinds (Xu, 2002, 2007; Xu & Carey, 1996). To summarize, Xu and Carey proposed a two-system account according to which 12-month-olds are sensitive to property/kind information that 10-month-olds are not because 12-month-olds possess sortal concepts, which support their suggestion of a kind-based object individuation system, that 10-month-olds do not possess.

This hypothesis, however, has been contradicted by accumulating evidence that non-linguistic (or partially linguistic) individuals, such as young infants and non-human animals could be sensitive to property/kind information. Some of these successes in human infants were attributed to the use of simplified tasks (Wilcox, 1999; Wilcox & Baillargeon, 1998a; Wilcox & Schweinle, 2002). For example, in event-monitoring tasks there is one event for the subject to watch and respond to, whereas in the original event-mapping task the subject watches one event (e.g., an occlusion event) but responds to another event (e.g., a non-occlusion event, such as the screen lowers to reveal the object(s) hidden behind). On the other hand, the task could be simplified by shortening the event sequence, for instance, using a single-trajectory event instead of an event in which the object(s) reversed back and forth (e.g., Krøjgaard, 2007; McCurry, Wilcox, & Woods, 2009; Wilcox & Schweinle, 2002). Both of these simplified tasks have revealed successful property/kind individuation at an earlier age (i.e., before the age of 10-month-old). Furthermore, there is evidence that younger infants can individuate objects based on ontological information, such as human, animal, or physical object, which is in direct conflict with the Carey-Xu account (Bonatti et al., 2005, 2002; Kibbe & Leslie, 2019; Surian & Caldi, 2010).

As explained in Chapter 1, to explain these mixed results, Stavans, Baillargeon, and collaborators proposed a new model involving the interaction between an object-file (OF) system and a physical reasoning (PR) system. This model (OF+PR model hereafter) suggests that young infants may already have both systems functioning for object individuation. The OF system encodes the spatiotemporal and categorical information, whereas the PR system encodes the event-related featural information. Critically, the disagreement between the two systems may lead to individuation failures, depending on the task. In short, only a quantitative disagreement (“2 versus 1”) after the event ends is unresolvable, as shown in Xu and Carey’s original property/kind task. On the contrary, this model predicts success in all the other paradigms discussed above: in the event-monitoring cases, since the subject responds during the ongoing event, the PR system has priority over the OF system to predict or interpret the outcome; in the single-trajectory task, the disagreement between systems is qualitative (OF: no object; PR: one object) and thus resolvable, since there is only one object that remains out of view; and in the last case, when the difference between objects is ontological, the OF system recognized the categorical difference whilst the PR system encodes the featural difference, so that both systems reach an agreement that two different objects are involved. In sum, this new model, as the authors claimed, seems to be able to explain all the current results in infants.

As discussed in Chapter 1, comparative studies also challenge the Carey-Xu account (Bräuer & Call, 2011; Fontanari, Rugani, Regolin, & Vallortigara, 2011; Kersken et al., 2020;

Mendes et al., 2008; Santos et al., 2002; Uller et al., 1997). Unlike the mixed results in infants, the nonhuman primate studies have shown both spatiotemporal and property/kind object individuation in the few species studied so far (Kersken et al., 2020; Mendes et al., 2008; Santos et al., 2002; Uller et al., 1997). These comparative studies didn't only challenge the strong version of the language gain hypothesis, but even implied that Xu and Carey's two representational system account may not be correct (Xu & Carey, 1996; Carey & Xu, 2001). The Carey-Xu account was derived from the empirical evidence of a dissociation between individuating by spatiotemporal and by property/kind information in human infants: 10-month-olds can do the former but not the latter. If we presume that ontogeny recapitulates phylogeny here, one piece of supportive evidence from comparative or evolutionary psychology would be that at least one species can be found to have and only have the object-file system to individuate objects using spatiotemporal information (like 10-month-old human infants), but not the second system based on property/kind information (like 12-month-old or older human infants). However, so far, no such dissociation between the two forms of object individuation has been found in non-human animals. All the adult nonhuman primates tested so far can do both. There is even evidence that domestic dogs and newly-hatched chicks (< 3 days old) succeed in similar tasks, suggesting they might be able to engage in some sort of spatiotemporal and property/kind individuation (Bräuer & Call, 2011; Fontanari et al., 2011, 2014). Since the performance of nonhuman primate infants has not been reported yet, we do not know if in these tests primates develop the property/kind individuation ability later than the spatiotemporal ability, like human infants. Nonetheless, we argue that the evidence of ontogenetic dissociations and that of phylogenetic dissociations may have very different implications. With ontogenetic dissociations, it is debatable that there is a single object representation module which supports initially spatiotemporal individuation, and then it also underlies property/kind individuation when there is enough prior knowledge, such as the knowledge about features, objects, physical principles, or linguistic knowledge (Baillargeon et al., 2012; Cacchione, Schaub, & Rakoczy, 2012; Carey & Xu, 2001; Stavans & Baillargeon, 2016; Wilcox et al., 2008). To some extent, we may take this as a moderate version of Carey-Xu's two-system account, or the one-system-plus account (see section 1.1.2).

On the contrary, phylogenetic dissociations can serve as a better evidence to support a radical version of Carey-Xu's two-system account: if spatiotemporal and property/kind individuation are based on two separate systems, it is plausible to find a species having one system while missing the other one, so that an adult individual of this species can pass only the spatiotemporal individuation task but not the property/kind task (or the other way around, hypothetically). To summarize, the possibility of an ontogenetic dissociation ("one can individuate objects first by spatiotemporal and later by property/kind information") in tested nonhuman primate species does not affect the relevance of addressing the lack of evidence of phylogenetic dissociations ("the adult individuals of a species can do one kind of individuation but not the other").

Hence, comparative findings insofar do not favour the Carey-Xu account. Xu and Carey's theory that the late-developing, second system of kind-based representation in human infants is driven by language would predict exclusively spatiotemporal individuation in nonhuman animals, but to repeat, so far there is no evidence in phylogeny that a species can use spatiotemporal information to individuate objects but not property/kind. However, it might be possible to try to explain these findings within the framework of the Carey-Xu account. One possibility is that there is in fact a second system to individuate objects using property and kind information, but this is evolutionarily older and independent of language, which would not be a necessary pre-requisite for property/kind individuation but just a facilitator in humans. The challenge would then be to find the species, or the key point in evolutionary history, corresponding to that developmental gap between 10-month-old and 12-month-old human infants. To examine this possibility, we need to investigate object individuation skills in further primate species.

The more recent and integrative OF+PR model does not make any specific predictions about the evolution of individuation abilities (Stavans et al., 2019). However, one would assume, based on the logic of the model, that it should be possible to find species that fail the sort of tasks that it considers to more difficult due to the un-resolvability of the conflict between systems. For example, the OF+PR model suggests that Xu and Carey's original property/kind task are more difficult, and this is why human infants fail it before the age of 12-months. One possibility is therefore that nonhuman primates also find this task more difficult. However, Uller et al.'s experiment using looking time measures and Santos et al.'s experiment using searching measures both showed that rhesus macaques individuated objects by property/kind information in the original paradigm (Santos et al., 2002; Uller et al., 1997).

Of course, such evidence is not necessarily a challenge to the OF+PR model, if we assume that the adult rhesus monkeys tested already reached the individuation level of 12-month-old human infants, who pass the original versions of both spatiotemporal and property/kind individuation task. However, even if we accept this explanation, one may still expect the possibility of finding species, for example more distant primate species, where the integration between systems fails in a similar way to 10-month-olds. More specifically, based on the OF+PR model, we would expect to find a species that fails the ‘difficult’ original version of the property/kind individuation task, but passes the “remainder” version, such as the box paradigm that most of primate research has used so far (Stavans et al., 2019). Indeed, if we analyse this box paradigm from the perspective of the OF+PR model, because the subject is required to determine whether there is any object remaining in the box after an object is removed, the problem is “qualitative” (“1 vs 0”, or “nothing vs something”), and therefore resolvable. This theory, therefore, would also benefit from comparative/evolutionary research, and more specifically from an extension of research on object individuation to more distantly related nonhuman primate species.

The first aim of the present study is to provide such an extension of research on spatiotemporal and property/kind individuation to new primate species in search of the potential phylogenetic dissociation between both systems that would correspond to the ontogenetic gap reported in human infants at 10 months.

2.1.2 Looking vs action response measurements

An additional aim of the present study is to combine in a single study the two types of measurement --looking time and action-based measurements in response to violation of expectations—that are normally used separately in object representation studies.

For example, Uller et al. used Xu and Carey’s original task with looking time measures and found that rhesus macaques were sensitive to property/kind information (as well as spatio-temporal information) (Uller et al., 1997). Later studies with free-ranging rhesus macaques measured 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 nonmatching pre-baited food, and also found successful individuation based on both types of information (Phillips & Santos, 2007; Santos et al., 2002). Mendes et al (2008) also used a box task and found that great apes (bonobos, chimpanzees and gorillas) were sensitive to both spatiotemporal and property/kind information, as they would continue to search inside the box even after the first retrieval if the object that they initially found was inconsistent with the input they saw. This finding was replicated in follow-up studies showing that great apes are sensitive to property/kind information, with either searching or begging behaviour measures (e.g., Bräuer & Call, 2011; Mendes et al., 2011).

The only study so far that has combined both types of measurements is Kersken et al. (2020), in which capuchin monkeys were tested with a similar box task including both looking and searching measures. Similar to the other box tasks, such as Mendes et al.'s experiment, the box in Kersken et al.'s also had a top opening where the food could be dropped into, and a secret floor panel that stopped the inserted food (Kersken et al., 2020; Mendes et al., 2008). Thus, the food item(s) revealed in the bottom compartment were always pre-baited food items, which may or may not be the same as the input (qualitatively or quantitatively). How a subject could access the food item depended on the panel installed on the bottom opening: the subject could access the food item only visually via a transparent plexiglass panel in the looking time trials, or they could access the items manually via an access hole in the manual search trials. The looking and searching measures yielded convergent results that capuchin monkeys not only use spatiotemporal information to individuate objects (looking or searching longer when the two objects were inserted sequentially but only one was found), but also use property/kind information to individuate objects (looking or searching longer when an object A was inserted but an object B was found) (Kersken et al., 2020).

The key question is whether these two measures are tracking the same ability. In general, this issue has been debated for several decades in different research domains, in infants and nonhuman animals, because in some research fields the performance revealed by different measures does not converge. For example, in several solidity tasks, human infants usually pass the looking task (Spelke et al., 1992), while older toddlers may fail the searching version of the task (Hood, Carey, & Prasada, 2000), when subjects were presented with a pair of events in which an object either stops at or moves through a solid barrier. This dissociation could not be simply explained by a lack of performance ability, as it also appeared in adult monkeys who had adequate performance ability for action (Cacchione et al., 2012; Hauser, 2001; Santos & Hauser, 2002). Moreover, the looking and searching performance of the same population can be uncorrelated. For instance, in an invisible object displacement task, younger toddlers passed the looking time version, but failed the searching version, and the two scores didn't correlate (Hood, Cole-Davies, & Dias, 2003).

In the field of object individuation, our previous experiment using the same design and including both looking and search measures showed that capuchin monkeys passed both versions of the spatiotemporal and property/kind individuation tasks as a group. However, at the individual level both measurements did not correlate (Kersken et al., 2020). One possibility is that variability in factors such as attention, motivation, or motor skill, rather than the individuation ability, prevented this correlation. Moreover, some of the measures did not allow much variability. For example, the reaching bout measurement in each trial typically varied between one or two reaches, which might make the detection of a correlation with the continuous measure of looking time difficult. Except Kersken et al.'s study, there is no other direct comparison across these two kinds of measure to our knowledge (Kersken et al., 2020).

Yet, there is evidence that measures of visual search and manual action converge in human infants, in terms of the age of emergence. Van der Walle et al. (2000) replicated Xu and Carey's (1996) looking time 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. Simplified versions of looking time and manual search tasks also found convergent results that younger infants use both spatiotemporal and property/kind information to individuate objects (McCurry et al., 2009; Wilcox & Baillargeon, 1998a; Wilcox & Chapa, 2004; Wilcox & Schweinle, 2002). Besides, macaques have shown competence in using both types of information to individuate objects in looking time tasks as well as the "searching time" used by Santos and colleagues, which combines (though does not distinguish) looking and manual search (Phillips & Santos, 2007; Santos et al., 2002; Uller et al., 1997).

In sum, existing evidence about the convergence of looking and search measurements is mixed: looking versus manual searching performance may not converge in some domains of knowledge, though a complete dissociation has not yet been found in the field of object individuation. Two interpretations have been proposed to explain the cases of dissociation in other domains. The first interpretation is that searching requires "stronger" or more robust representations than looking (Munakata, 2001). For instance, tasks measuring looking time may only require the subject to detect a violation (post-hoc or "post-diction"). To do so, a subject does not necessarily need to understand the violation and make accurate predictions, which is required for planning and conducting an action (Cacchione & Burkart, 2012; Hood et al., 2003). The dissociation between looking time and manual searching tasks may also reflect the existence of multiple systems, such as a "what" system for object recognizing, and a "where" system for object-directed grasping (Goodale & Milner, 1992; James, Culham, Humphrey, Milner, & Goodale, 2003; Leslie et al., 1998; Milner & Goodale, 2008).

In contrast, a second explanation suggests that the dissociation between measures may not reflect the existence of multiple different systems, but simply a difference in terms of task demands or the power to reveal the competence. It is suggested that the looking task is a more sensitive measure, since it is less likely to be confounded with some other skills, such as planning, memorizing, or inhibiting (Berthier et al., 2001; Hood et al., 2003; Mash, Novak, Berthier, & Keen, 2006; Santos & Hauser, 2002; Santos, Seelig, & Hauser, 2006). Although it is still unclear how best to explain cases of divergence between looking time and manual search measures, we decided to include both kinds of measures to obtain a detailed picture of the object individuation ability in our species, and to allow a direct comparison with our previous findings in capuchin monkeys. The looking time measure may be useful to detect a weaker representation or competence that might be masked due to confounding factors, whereas the manual searching measures, as a more stringent one, may provide a stronger evidence of full competence.

2.1.3 Aims of current study

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, squirrel monkeys with a box task previous used with capuchin monkey to investigate their property/kind and spatiotemporal individuation skills (Mendes et al., 2008; Santos et al., 2002; Kerskens, et al., 2020). This is the second New World monkey species to be tested with this paradigm, after its sister species, the capuchin monkey (Kersken et al., 2020).

Capuchin monkeys are known for their relatively good competence in relation to object representation (Mathieu et al., 1976; Schino et al., 1990) and even object manipulation or tool-use (Fujita et al., 2003; Luncz et al., 2016; Moura & Lee, 2004; Visalberghi, Fragaszy, & Savage-Rumbaugh, 1995). Squirrel monkeys, in contrast to capuchin monkeys, although sharing the same ecological environment, have been so far reported to be less skilful in terms of object cognition (de Blois et al., 1998; Jaakkola, 2014). We did not have clear predictions about how squirrel monkeys would perform. They might show competence in both individuation tasks using either type of information, as the other nonhuman species tested to date. Or, if property/kind object individuation is cognitively different to spatiotemporal individuation as the Carey-Xu account suggested, we might be able to find a dissociation in this species similar to that reported in 10-month-old infants in the traditional tasks: squirrel monkeys might individuate objects based on only spatiotemporal information, but not property/kind information (Carey & Xu, 2001).

Secondly, we aimed to measure both looking time and manual searching behaviours to explore the relationship between the two measures at a group level (whether the subjects pass or fail the tasks with both measures) and at an individual level (whether the subjects' performance on either measure correlated or not). If the ability to individuate objects based on spatiotemporal versus property/kind information relies on different systems or cognitive abilities, we might find differential dissociations in each type of tasks (for example, convergently passing the spatiotemporal task with looking and reaching measures, but only passing the looking time version of property/kind task). Parallel dissociation across both types of individuation would also make a strong case for the Carey-Xu account (e.g., passing both looking and searching versions of spatiotemporal tasks but failing both versions of the property/kind condition). Otherwise, the convergence in both conditions using both measures would be more difficult to explain under Xu and Carey's account (Carey & Xu, 2001; Xu & Carey, 1996). Although our study was designed and completed before the publication of the new theoretical framework proposed by Stavans et al., and although this model does not make any specific evolutionary or comparative predictions, we will discuss the potential implications of our results for this approach (Stavans et al., 2019).

2.2 Method

2.2.1 Subjects

Eighteen captive female common squirrel monkeys (*Saimiri sciureus*) from two social groups (East and West groups at the Living Links Research Centre at Edinburgh Zoo) participated in the present experiment. The monkeys were divided into two matched experimental groups and then took part in looking time and manual search tasks in counterbalanced orders: half of the monkeys completed the looking time trials first (looking-first group hereafter) and half of them completed the manual search trials first (searching-first group hereafter). The demographic information of each group can be found in Table 2-1.

All monkeys were born in captivity, raised by their mothers, and live at the Living Links Research Centre at Edinburgh Zoo (for more details about the facility: Macdonald & Whiten, 2011). All monkeys have previously taken part in behavioural studies at the site. They lived in two groups with similar size, identical provisions and enclosures. Each group' enclosure contains an outdoor part as well as an indoor part, which is connected to the cubicles in the experiment room via slide doors and tunnels. These entrances for the squirrel monkeys were only open during the experimental sessions (2 sessions maximum per day, each sustains about 1.5 hours) so that the monkeys can access the cubicles on a voluntary basis. The monkeys were not food or water deprived and they were free to leave during the studies. Only positive reinforcements (food) were used. The study was granted ethical approval and supervised by the University Teaching and Research Ethics Committee of University of St Andrews.

Table 2-1 The demographical information and food preference of subjects

Name	Experiment group	Natural group	Age (yrs)	Food preference		
				Peanut	Raisin	Binomial test (p)
Roca	LT	East	14	5	5	>0.99
Cali	LT	East	12	6	4	0.75
Lexi	LT	East	7	7	3	0.34
Amarilla	LT	East	6	1	9	0.02
Gabriela	LT	East	6	3	7	0.4
Valencia	LT	East	6	2	8	0.11
Elie	MS	East	12	9	1	0.02
Maya	MS	East	14	10	0	0.01
Sipi	MS	East	8	5	5	>0.99
Flora	MS	East	9	4	6	0.75
Dora	MS	East	7	4	6	0.75
Pelusa	MS	East	6	4	6	0.75
Ciara	MS	East	6	3	7	0.34
Sancha	LT	West	8	4	6	0.75
Orla	LT	West	6	4	6	0.75
Loki	LT	West	3	5	5	>0.99
Jasmin	MS	West	16	7	3	0.34
Toomi	MS	West	11	8	2	0.11
Gisele	MS	West	5	5	5	>0.99

Note: LT refers to "looking-time-first group" whereas MS means "manual-search-first" group

2.2.2 Apparatus

The main apparatus of the present experiment was an opaque plastic box (see Figure 2-1 as a sketch). It has (a) a circular opening on the top for the experimenter to drop food into the box, (b) a lower front opening with a set of Plexiglas sliders (a black opaque one, one with a rounded rectangular access hole for the manual search trials, and a transparent one for the looking time trials) which enable the participants to look/reach into the box depending on the condition, (c) a flap at the lower back of the box through which the experimenter can preload / retrieve the reward in the bottom compartment, and most importantly, (d) a slider on the secret second floor (invisible from the participants'' point of view) to hold the dropping food in the test phase, which can also be removed to allow the monkey to see the food items dropping to the ground in the familiarization phase. Carpet tiles and towels was used in the lower and upper compartment respectively to eliminate the potential sound cues caused by dropping.

The monkeys were introduced into and then separated in the cubicle with a window to the front containing an elongated hole, through which the monkeys could extend their arms towards the box placed on a trolley table (see Figure 2-2). At the start of the trial, the box was roughly 40 centimetres away from the cubicle, with the front opening of it (especially, the access hole on the slider in manual search trials) aligned to the hole on the cubicle, so that the monkeys could look or reach into the box without too much effort when the box was pushed towards the cubicle, and the sliders were open.

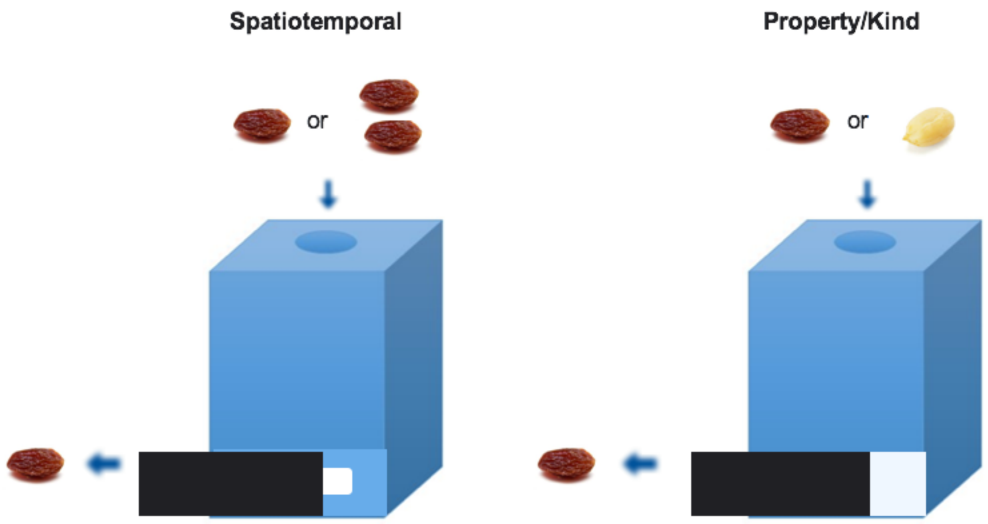


Figure 2-1 Schematic representation of the apparatus and stimuli (the left panel present the slider with access hole for manual search conditions; the right panel shows the transparent slider for looking time conditions).

GoPro cameras controlled with a GoPro smart remote were used to record the entire scene. A GoPro Hero 4 was fastened to the front of the box right above the front opening to enable the coding of the looking behaviours. A GoPro Session 4 was attached to the inner back wall of the box so that it filmed the monkeys' reaching behaviours from inside the box. An additional camera was placed on a tripod to record the whole view of the box and the experimenter so that we could check for potential procedural errors or environmental distractors. Besides, in the looking time trials a LED light was attached on the ceiling of the main compartment to shed light on the food reward, ensuring that the monkeys could see it clearly.

Food rewards included sunflower seeds, raisins and shelled peanuts. Monkeys were rewarded with seeds for coming into the cubicle and for participating in the experiment at the beginning and the end of each trial, respectively (hereafter, cubicle rewards). The raisins and peanuts were used as stimuli in the test, after preference testing revealed that these two foods were equally preferred by the monkeys as a group (see below). The preloaded food or the food in the outcome always was the kind that the individual monkeys preferred, depending on a preceding food preference test, to make the experience as similar as possible across subjects. The type and amount of food rewards used in the experiment was under the advice of the keepers at Living Links.

Figure 2-2 The apparatus and set-ups from the view of each cameras (left-top: front view of the monkey and the cubicle panel; left-bottom: top view from inside the box; right: side view of the whole set-ups before experiment)

2.2.3 Procedure

Food preference test

Before the experimental task, we conducted a food preference test for all the participants to choose between two kinds of food with equivalent value. Monkeys had to choose one of two different food rewards (chosen among sunflower seed, shell peanut, raisins, diced dates, and maize) presented on a piece of carpet. We repeated each pair of contrasts for 2 to 4 times and found raisin versus peanut was the most feasible pair in consideration that these two kinds of food are of similar size, both easy to manipulate, and roughly equally preferred. The monkeys then took part in a second-round food choice task just between peanuts and raisins for ten trials. The results showed these two kind of reward were equally attractive to squirrel monkeys at a group level, with only three out of eighteen monkeys showing a significant preference. Hence, we used raisins and peanuts for the test and only used the preferred food for each individual monkey as the preloaded reward.

Familiarization phase

Before the test phase, each subject completed several familiarization trials to ensure the monkeys got used to the apparatus and understood the set-up. The procedure for the familiarization trials is described below. Note that the familiarization trials were almost the same between the two experimental groups (searching first vs. looking first) except for 2 small differences: (1) the monkeys in the searching group were allowed to reach for the food after the demonstration, whereas the monkeys in the looking first group could only watch and get the food from the experimenter after roughly 15 seconds; (2) an object retrieval trial was only provided for monkeys in the searching-first group. All monkeys went through the familiarization phase only once to avoid habituation. So, there was no second round of familiarization for the monkeys when switching measures.

Familiarization for Object retrieval task (only for searching-first group): Subjects were presented with the box and then allowed to reach into the bottom compartment to retrieve the food reward after the opaque slider had been removed and the trolley moved forward. Each monkey in the searching first group needed to succeed only once in this task to proceed to the next one. In case some of the monkeys never tried to approach the apparatus because of neophobia, we also prepared a scaffolding task in which the food was placed in front of the box (on the table) so that the monkeys don't have to put their hands into the novel apparatus at the very beginning. A monkey who failed in the first attempt of object retrieval had to pass the scaffolding task before their second chance.

Transparent trial: Each monkey was presented with the box while both the opaque slider and the secret floor were removed, so that they had full visual access to the bottom compartment via the transparent slider. A food reward was placed beside the box to attract the monkey's attention, and was then dropped into the box. After watching the food dropping onto the floor of the compartment, the box was moving forward (moving straight to the cubicle and transparent slider removed for the manual search group but only moving slightly forward for the looking first group) to allow the monkey to access the bottom compartment manually or visually.

Opaque trial: The procedure was the same as in the transparent trial except that the opaque slider was inserted before the trial and removed after the dropping so that the monkeys couldn't see the food land on the bottom. The monkeys were allowed to access the bottom compartment manually or visually after the opaque slider was removed.

Experimental phase

The design was adapted from Santos et al. (2002). The trials shared the same basic procedure except for subtle differences in the food dropped into the box (see Table 2-2): in the spatiotemporal conditions, the number of food items inserted was manipulated; in the property/kind conditions, the kind of food inserted was manipulated. In each condition the outcome was always the monkey's preferred food, so that the reaction could not be simply explained by the immediately perceived information. The food preloaded in the box that the monkey would later find at the end of the trial, was either the same as the food inserted (consistent trials) or not (inconsistent trials). Each monkey went through 8 trials (4 conditions across 2 measures) in total. The looking time and manual search measures were counterbalanced across blocks: half of the monkeys completed the looking time block first and the half of them completed the manual search block first (see Table 2-1). The order of the four conditions within each measure blocks was counterbalanced.

Table 2-2 Baited and retrieved food items for each condition.

Condition	Food inserted	Food outcome	Additional food
Property/Kind Consistent (PKC)	A	A	-
Property/Kind Inconsistent (PKI)	B	A	B
Spatiotemporal Consistent (STC)	A	A	-
Spatiotemporal Inconsistent (STI)	AA	A	A

Note: A and B denote different types of food (either half peanuts or raisins);A denotes the more preferred food for the particular subject. "Additional food" refers to the "missing" food item in the inconsistent trials, which was later retrieved and rewarded to the monkeys at the end of the trial.

Before each trial, the box was placed on the trolley about forty centimetres away from the cubicle with both sliders inserted (the second floor, the opaque slider on top of the transparent one or the one with the access hole). A food reward was preloaded secretly in the main compartment of the box. After these preparations, a monkey was introduced in the separated cubicle and rewarded with a seed. When the monkey finished eating, the experimenter attracted its attention by presenting the experimental food item(s) beside the opening of the box, and then lifted and dropped the food into the box while the monkey was looking (one by one in the spatiotemporal inconsistent trials). The box was then pushed forward to the monkey after the opaque slider was removed.

For the manual search trials, the box was slid straight to the cubicle window to enable the monkey to access it for 35 seconds. After the searching period, the box was pulled back with the opaque slider inserted. The experimenter opened the back flap and pretended to search the compartment. In the inconsistent trials, the additional food reward was then handed to the monkey as if it was stuck somewhere or stayed on a blind corner.

For the looking time trials, the box was moved slightly forward staying roughly 20 centimetres away from the cubicle. The monkey was allowed to watch for 25 seconds. After that, the experimenter inserted the opaque slider, pulled back the box, searched the compartment, and took the food that had been contemplated to the monkey. In the inconsistent trials, the experimenter would rummage the box to pretend he was retrieving the additional food to the monkey as in the manual search inconsistent trails.

At the end of the trial, the monkey received another seed whilst the experimenter showed the end signal (showing the monkeys both hands wide open) and opened the cubicle slider for the monkey to leave.

Coding

The program BORIS for Mac (Friard & Gamba, 2016) was used to code the videos. Each video was first cropped into (1) an attentional check section including the demonstration of dropping and (2) a coding period (25 seconds for looking time trials; 35 seconds for manual search trials) starting at the time point that the opaque slider was removed. Both sections were muted and renamed so that the coder was blind to the information of conditions.

Looking time trials

The monkeys' total looking time and the number of looking bouts were coded as dependent variables for the looking time trials. The monkeys' total looking time was defined by the total looking duration that the monkey watched the lower part of the box and looked directly at the content during the coding period (twenty-five seconds since the removal of the opaque panel). The looking bout was counted frame-by-frame (30Hz or 33.3 milliseconds per frame). A looking bout was coded when a subject looked into the box for two or more consecutive frames (> 66.6 ms) before it clearly looked away for at least one frame.

Manual search trials

The monkeys' total reaching time and the number of reaches were coded as the main dependent variables for the manual search trials. The coding was primarily based on the footage from the inside-box camera. During the thirty-five seconds period since the removal of the opaque panel, the duration of all the instances of reaching through the access hole was calculated as the total reaching time using the same method as above, whilst the total number of reaches was counted frame-by-frame as another index. Similar to the looking time measures, reaching duration was counted frame-by-frame, whereas a reaching bout was coded when a subject reached into the box for two or more consecutive frames before it took the hand(s) out.

In addition to the total reaching duration and the total number of reaching bouts, we also coded the "second search", defined as reaching after retrieval. Namely, after the monkey took the food item (and its hands) out of the box, the following reaching behaviours (if there were) were coded as "second search", including the second reaching duration and the second reaching bout, coded in a way similar to the looking and the total reaching measures.

2.3 Results

A number of trials had to be excluded from the analysis or to be redone (see Table 2-3) because of experimental errors, such as food reward rolling out of sight or out of reach, disruption in the environment, or the connection of the GoPro camera failing during recording. If the monkeys did not react or left the experiment early, or any disruptions above happened before the dropping event, the trial was redone.

Table 2-3 The total number of analysed and excluded trials for each condition.

Measures and conditions	No. of trials for analysis	No. of excluded trials	No. of redone trials
Looking Time	(n=19)		
STC	19	0	0
STI	18	1 (no reaction)	0
PKC	19	1 (error)	1
PKI	19	0	0
Manual Search	(n=18)		
STC	18	0	0
STI	17	4 (1 error, 3 no reaction)	3
PKC	17	4 (3 errors, 1 no reaction)	3
PKI	17	3 (1 error, 2 no reaction)	2

Note: For measures and conditions, "ST/PK" means spatiotemporal/property/kind, "C/I" means consistent/inconsistent. For the exclusion reasons, "error" means the camera failed, "no reaction" means the subject stopped participating in the middle of the events. Most but not all these trials were redone (trials were not redone in the case that a monkey watched the event to the end). The number of trials for analysis is the total number of trials (i.e., the number of participants) minus the number of excluded trials plus the number of trials redone.

2.3.1 Looking time measure

Considering that a major part of the data was not normally distributed, we used a Wilcoxon Signed Ranks test to compare the median looking and reaching bouts and duration across the conditions, in addition to a two-way ANOVAs with Task and Consistency as factors for each measure.

We calculated the monkeys' mean number of looks, and the total looking time directed at the target for each condition (Figure 2-3).

Spatiotemporal Condition

Monkeys didn't look significantly longer at the inconsistent events than at the consistent events in spatiotemporal trials ($F(1,15) = 0.775$, $p = 0.392$, $\eta^2 = 0.049$). No difference was found between looking-first and searching-first groups ($F(1,15) = 0.045$, $p = 0.835$, $\eta^2 = 0.003$). The interaction between Group and Consistency was also not significant ($F(1,15) < 0.001$, $p = 0.991$, $\eta^2 < 0.001$).

Monkeys didn't look significantly more times at the inconsistent events than at the consistent events in spatiotemporal trials ($F(1,15) = 0.033$, $p = 0.859$, $\eta^2 = 0.002$). No difference was found between looking-first and searching-first groups ($F(1,15) = 0.995$, $p = 0.334$, $\eta^2 = 0.062$). The interaction between Group and Consistency was also not significant ($F(1,15) = 0.033$, $p = 0.859$, $\eta^2 = 0.002$).

Property/kind Condition

Monkeys didn't look significantly longer at the inconsistent events than at the consistent events in spatiotemporal trials ($F(1,16) = 1.476$, $p = 0.242$, $\eta^2 = 0.084$). No difference was found between looking-first and searching-first groups ($F(1,16) = 0.099$, $p = 0.757$, $\eta^2 = 0.006$). The interaction between Group and Consistency was also not significant ($F(1,16) = 0.027$, $p = 0.870$, $\eta^2 = 0.002$).

Monkeys didn't look significantly more times at the inconsistent events than at the consistent events in spatiotemporal trials ($F(1,16) = 0.197$, $p = 0.663$, $\eta^2 = 0.012$). No difference was found between looking-first and searching-first groups ($F(1,16) = 1.343$, $p = 0.264$, $\eta^2 = 0.077$). The interaction between Group and Consistency was also not significant ($F(1,16) = 0.018$, $p = 0.894$, $\eta^2 = 0.001$).

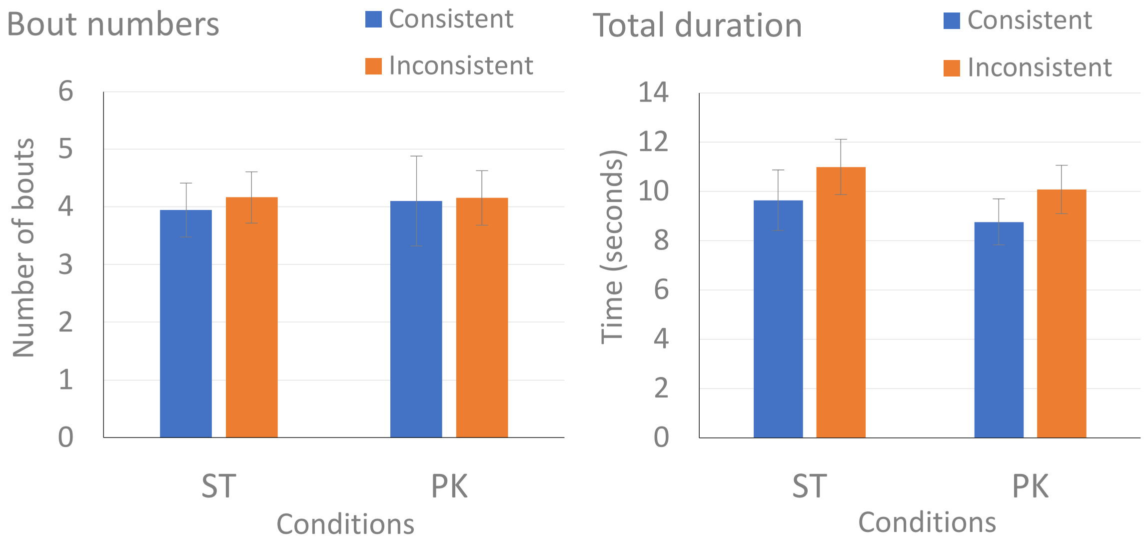


Figure 2-3 Squirrel monkeys' looking behaviours in manual search trials. Left panel: looking bouts; right panel: looking duration.

2.3.2 Manual Search measures

We calculated monkeys' mean reaching time and number of reaches into the box for each condition (Figure 2-4).

Spatiotemporal Condition

Monkeys didn't reach in significantly longer at the inconsistent events than at the consistent events in spatiotemporal trials ($F(1,16) = 1.605$, $p = 0.223$, $\eta^2 = 0.091$). No difference was found between looking-first and searching-first groups ($F(1,16) = 0.683$, $p = 0.421$, $\eta^2 = 0.041$). The interaction between Group and Consistency was also non-significant ($F(1,16) = 2.329$, $p = 0.146$, $\eta^2 = 0.127$).

Monkeys didn't reach in significantly more often at the inconsistent events than at the consistent events in spatiotemporal trials ($F(1,16) = 1.622$, $p = 0.221$, $\eta^2 = 0.092$). No difference was found between looking-first and searching-first groups ($F(1,16) = 0.732$, $p = 0.405$, $\eta^2 = 0.044$). The interaction between Group and Consistency was also non-significant ($F(1,16) = 1.136$, $p = 0.302$, $\eta^2 = 0.066$).

Property/kind Condition

Monkeys didn't look significantly longer at the inconsistent events than at the consistent events in spatiotemporal trials ($F(1,16) = 0.213$, $p = 0.651$, $\eta^2 = 0.013$). No difference was found between looking-first and searching-first groups ($F(1,16) = 0.522$, $p = 0.481$, $\eta^2 = 0.032$). The interaction between Group and Consistency was also not significant ($F(1,16) = 0.177$, $p = 0.681$, $\eta^2 = 0.011$).

Monkeys didn't look significantly more often at the inconsistent events than at the consistent events in spatiotemporal trials ($F(1,16) = 0.979$, $p = 0.337$, $\eta^2 = 0.058$). No difference was found between looking-first and searching-first groups ($F(1,16) = 0.410$, $p = 0.531$, $\eta^2 = 0.025$). The interaction between Group and Consistency was also not significant ($F(1,16) = 0.180$, $p = 0.677$, $\eta^2 = 0.011$).

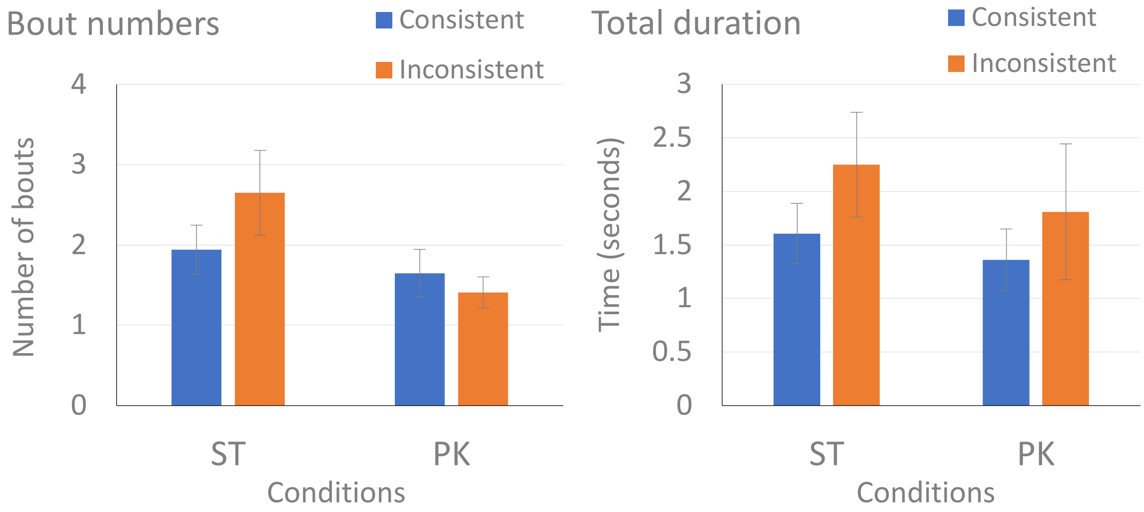


Figure 2-4 Squirrel monkeys' total reaching behaviours in manual search trials. Left panel: reaching bouts; right panel: reaching duration.

Second search measures

We also analysed the second search behaviours, namely any additional searching after retrieving the food items in the manual search tasks, calculating the duration and number of bouts for second searches (Figure 2-5).

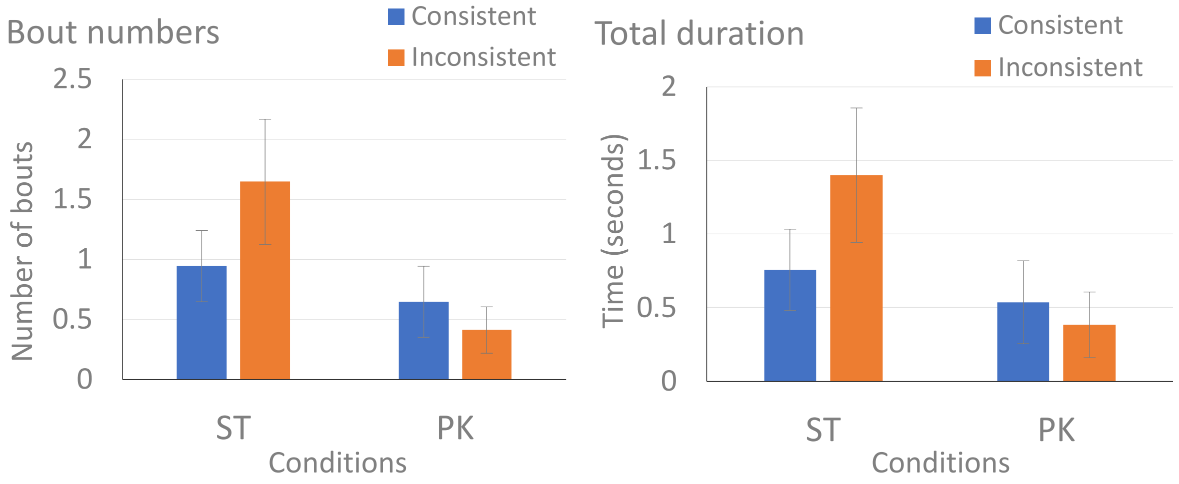


Figure 2-5 Squirrel monkeys’ second reaching behaviours (reaching after retrieval) in manual search trials. Left panel: reaching bouts; right panel: reaching duration.

Spatiotemporal Condition

Monkeys didn’t reach significantly longer in the inconsistent events than in the consistent events in spatiotemporal trials ($F(1,15) = 1.559$, $p = 0.231$, $\eta^2 = 0.094$). No difference was found between looking-first and searching-first groups ($F(1,15) = 1.029$, $p = 0.326$, $\eta^2 = 0.064$). The interaction between Group and Consistency was also not significant ($F(1,15) = 1.339$, $p = 0.255$, $\eta^2 = 0.085$).

Monkeys didn’t reach significantly more often in the inconsistent events than in the consistent events in spatiotemporal trials ($F(1,15) = 1.844$, $p = 0.195$, $\eta^2 = 0.109$). No difference was found between looking-first and searching-first groups ($F(1,15) = 1.407$, $p = 0.254$, $\eta^2 = 0.086$). The interaction between Group and Consistency was also not significant ($F(1,16) = 1.350$, $p = 0.263$, $\eta^2 = 0.083$).

Property/kind Condition

Monkeys didn't look significantly longer at the inconsistent events than at the consistent events in spatiotemporal trials ($F(1,14) = 0.177$, $p = 0.681$, $\eta^2 = 0.012$). No difference was found between looking-first and searching-first groups ($F(1,14) = 0.131$, $p = 0.723$, $\eta^2 = 0.009$). The interaction between Group and Consistency was also not significant ($F(1,14) = 0.949$, $p = 0.374$, $\eta^2 = 0.063$).

Monkeys didn't look significantly more often at the inconsistent events than at the consistent events in spatiotemporal trials ($F(1,14) = 0.412$, $p = 0.531$, $\eta^2 = 0.029$). No difference was found between looking-first and searching-first groups ($F(1,14) = 0.111$, $p = 0.744$, $\eta^2 = 0.008$). The interaction between Group and Consistency was also not significant ($F(1,14) = 0.412$, $p = 0.531$, $\eta^2 = 0.029$).

2.3.3 Correlation between Looking and Searching measures

There was no correlation between each pair of dependent variables of looking and searching behaviours. Looking duration was not significantly correlated with reaching duration (Spatiotemporal: $r(17) = 0.237$, $p = 0.359$; Property/Kind: $r(16) = -0.312$, $p = 0.239$). Number of looking bouts was not significantly correlated with the number of reaching bouts (Spatiotemporal: $r(17) = 0.205$, $p = 0.430$; Property/Kind: $r(17) = -0.212$, $p = 0.430$).

Apart from the frequencies and duration of behaviours in general, we also calculated the difference scores between consistent and inconsistent conditions in each task (i.e., difference = inconsistent – consistent) to approximately represent the ability of the monkeys to detect the violation. Again, no significant correlation was found for Spatiotemporal tasks (Number of bouts: $r(17) = 0.205$, $p = 0.430$; Duration: $r(18) = 0.244$, $p = 0.346$) nor for Property/Kind tasks (Number of bouts: $r(18) = -0.234$, $p = 0.351$; Duration: $r(18) = -0.355$, $p = 0.148$).

2.4 Discussion

Squirrel monkeys in the present study didn't pass either the spatiotemporal or the property/kind object individuation tasks, which is in contradiction to all the previous studies on other primate species reported so far, including the one by Kersken et al (2020) using exactly the same apparatus and procedure in the same setting. According to the existing literature using the same paradigm and similar tasks, the skill of individuating objects by spatiotemporal and property/kind information is present in apes, rhesus macaques, capuchins, and dogs, as well as in human infants older than twelve-months (Bräuer & Call, 2011; Fontanari et al., 2011, 2014; Kersken et al., 2020; Mendes et al., 2008; Van De Walle, et al., 2000; Xu & Baker, 2005). Note that young infants who failed to individuate objects by property/kind still succeeded in spatiotemporal object individuation tasks (Xu & Carey, 1996), suggesting the latter may rely on a more fundamental ability that younger infants already mastered by the age of ten months. In the current experiment, however, not only did the squirrel monkeys fail the property/kind task, but they also failed the putatively "easier" spatiotemporal task, making of them the first animal species reported to have failed an object individuation test.

Moreover, our study utilized both looking time and manual search measures in the same subject group with the same experimental apparatus, and we found that results from both measures converged at the group level: as a group, monkeys neither looked longer nor searched more often when the object they found at the bottom of the box was not the same object they saw going in, and when they found only one object after seeing two placed inside. At the individual level, the looking and search performances were not correlated with each other, i.e., the monkeys that looked longer or more often in the visual task did not tend to reach more in the manual search task. Note that this lack of correlation was also found in the capuchin object individuation study where the monkeys demonstrated both spatiotemporal and property individuation at the group level in both measures. This may reflect some noise factors such as motivation or attention, or the fact that the variation and variability allowed by the two different measures is very different and make the finding of a correlation difficult (Kersken et al., 2020).

In summary, our findings seem to suggest that squirrel monkeys do not have the ability to individuate objects by either spatiotemporal or property/kind information. If this interpretation is correct, a plausible and parsimonious explanation is that object individuation is not a primate primitive ability, and its presence in other primate and non-primate species should be explained by convergent evolution. An alternative possibility is that the common ancestor of modern primates[1] (including the squirrel monkeys, capuchin monkeys, rhesus monkeys, apes, and human beings) was equipped with the object individuation ability, but some species, such as the squirrel monkey, lost such ability during evolution. Either of these interpretations does not support or contradict Xu and Carey's two-system account given that squirrel monkeys failed both individuation tasks. According to our findings, squirrel monkeys are not the key species that would reveal the dissociation between spatiotemporal competence accompanied by property/kind incompetence, that would mimic the 10-month dissociation found in human infants, and that would be necessary to support Xu and Carey's hypothesis from an evolutionary perspective.

From this point of view, the absence of both skills in squirrel monkeys, along with the previous findings of co-occurrence of both skills in other nonhuman primates, would if anything cast doubt on the idea that the two kinds of object individuations rely on different representational systems (Carey & Xu, 2001; Xu, 2007; Xu & Carey, 1996). Spatiotemporal and property/kind individuation may share at least some basic cognitive underpinnings and they are either both present or both absent.

What are the implications of our findings for the recent Stavans and Baillargeon's OF+PR model (Stavans et al., 2019)? This model proposed the object-file (OF) and physical reasoning systems (PR) work in parallel to individuate objects by spatiotemporal and featural information respectively, and whether the individuation is successful or not depends upon whether the two systems agree or a conflict between systems can be resolved, which is highly task dependent. Specifically, the box task we used in the current chapter belongs to a type of task in which the potential disagreement between systems can be resolved, so that human infants younger than 10-month of age can pass. Thus, from the perspective of the OF+PR

---

[1] Given the evidence that dogs and chicks have shown object individuation ability, it is possible to find a more ancient origin for such ability. To examine this possibility and to extend our knowledge along this direction, further research may be needed to include more mammal and other species.

model, one may have predicted that squirrel monkeys should at least behave like 10-month-old human infants and be able to individuate objects by either type of information in the magic box task that we used , just as the previous research using similar tasks have revealed in apes, macaques, and especially capuchin monkeys (e.g., Kersken et al., 2020). However, our results with the squirrel monkeys are inconsistent with such prediction and with the previous empirical findings in closely related species.

However, although, our tentative conclusion from the results of this experiment should be that squirrel monkeys lack both spatiotemporal and object individuation abilities, there is also an alternative possibility, namely, that our task, even if identical to that used with capuchin monkeys in exactly the same setting and with the exact same procedure, may have failed to uncover an existing individuation competence in squirrel monkeys because there is something in it that makes it specifically unsuitable for this monkey species.

First, it is possible that the squirrel monkeys tend to react more “intensely”, for example by looking longer at the outcomes and searching more for the food. This may lead to a “ceiling effect” in both measurements that could mask any performance difference between conditions. Thus, no matter if the outcome was consistent or inconsistent with the input, the squirrel monkeys tended to look generally much longer (about ten seconds) at the food items in the bottom compartment of the box compared to the capuchin monkeys in the Kersken et al (2020) study (less than four seconds). A similar trend was found regarding the number of reaching bouts, with squirrel monkeys tending to search more overall than capuchin monkeys. These comparisons, though not statistically tested since the food targets were different, suggest that squirrel monkeys may be “opportunistic” in their approach to food procurement, i.e., that they may act first, likely redundantly, to best secure every potential food item, before or even without thinking or attending to the context of the situation.

The second possibility is that the present paradigm was not suitable for the physical characteristics of squirrel monkeys. The box was originally made for use with the capuchin monkeys, a larger species (i.e., a capuchin monkey can be about 40% larger in terms of height/length than a squirrel monkey), so that it might have been too large for squirrel monkeys. Thus, the increased relative distance between the upper part where the food was dropped, and the lower part where it had to be searched for may have made keeping track of the events more challenging for squirrel monkeys, that would need to shift their attention across relatively longer spatial areas. Moreover, the moving trajectories of the food items were complex (raised up from the table before being dropped down through the box), making it potentially more challenging for the monkeys to attend and follow the food items across that relatively extended space.

Last but not least, in this paradigm and apparatus there is only one searching area, the box opening at the bottom, making the searching behaviours ambiguous to interpret. For example, a second reaching attempt can be due to either (i) an individual believes there shall be a certain food item, or (ii) curiosity or general interest because of a higher arousal level after the first retrieval of food. In other words, a single-location paradigm with the box may lack the power to assess the object individuation ability in some species.

To test these two assumptions and further examine the squirrel monkeys' competence on object individuation, we conducted two additional experiments. In chapter 3, we compared squirrel and capuchin monkeys' reaction patterns towards the same stimuli moving in the same way using a computer-based touchscreen task. In chapter 4 and 5, we adapted Xu and Carey's two-location paradigm to test squirrel monkeys for their spatiotemporal and property/kind object individuation in a scenario that was better scaled to their body proportions and allowed for additional measurements taking advantage of the two locations (Xu & Carey, 1996).

Chapter 3 Comparing squirrel and capuchin monkeys' reaction pattern: "Whack-a-cricket" experiment

3.1 Introduction

The squirrel monkeys didn't succeed on either the spatiotemporal or the property/kind object individuation tasks in Chapter 2, unlike a wide range of primate species including human infants before one-year-old, great apes, macaques, and capuchin monkeys (Kersken et al., 2020; Mendes et al., 2008; Santos et al., 2002; Stavans et al., 2019). This was, for example, the case of capuchin monkeys, a close relative of squirrel monkeys that in the wild share the same ecological environment (as well as in the Living Links Centre where these species share the same facility). A previous study using exactly the same paradigm and apparatus with the Living Links capuchins revealed that capuchin monkeys generally searched only once in one-object trials, while they searched averagely about twice when the spatiotemporal or property/kind information suggested there were two objects (Kersken et al., 2020). This finding suggests that the capuchin monkeys individuated the object(s), correctly determined the number of objects involved in the scenario, and used these object representations to guide their behaviours. In contrast, the squirrel monkey in the experiment presented in Chapter 2 were likely to search more than once even when there should be only one food item in the trial.

What does this mean? A simple explanation for squirrel monkeys' failure would be the lack of ability to individuate objects based on either spatiotemporal or property/kind information. Nonetheless, apart from this "true failure" account, it is also possible that the squirrel monkeys are able to individuate objects but could not express their competence in this particular task. Two possibilities were mentioned in the last chapter. One is that the squirrel monkeys didn't form or use the object representation because in a foraging context like the magic box, they "act" as fast as possible, searching for the food before they "think" properly (processing the information to represent the objects), while the capuchins (and probably other primates we mentioned above) are more "deliberate".

Hence, we propose a “Cursory squirrel monkey vs. deliberate capuchin monkey” hypothesis (“hypothesis” hereafter in this chapter). This performance difference could possibly be explained by a trade-off between reaction speed and the accuracy and precision of actions. Some empirical findings about squirrel monkeys seem to be in line with this hypothesis. In general, it is reported that squirrel monkey seems to be short at maintaining attention and self-control (Anderson, Kuroshima, & Fujita, 2010; Fragaszy, 1985; MacLean et al., 2014). Squirrel monkeys performed poorly on inhibitory control tasks compared to capuchin monkeys (“A not B” and Cylinder tasks, MacLean et al., 2014). In a delayed response task, squirrel monkeys acted hastily to respond towards a positive locus (the location with food), or even a negative locus (the location without food) when the positive one was temporarily inaccessible (French, 1959). There is evidence that after some facilitating intervention was applied, squirrel monkeys could successfully wait for food rewards to accumulate as well as capuchin monkeys (Anderson et al., 2010).

Due to the scarcity of research, it is hard to draw firm conclusions about the extent to which squirrel monkeys are able or willing to “wait” in contrast to capuchins and other primate species. In contrast, there are many more such studies on capuchin monkeys, suggesting a) relatively good delay tolerance (Addessi et al., 2013; Addessi, Paglieri, & Focaroli, 2011; Amici, Cacchione, & Bueno-Guerra, 2017; Evans, Beran, Paglieri, & Addessi, 2012). For instance, when choosing between having immediate access to two pieces of food or delayed access to six pieces of food, capuchins can wait for the latter option. Intriguingly, the period of time capuchins was willing to wait is longer than tamarins and marmoset monkeys, and were quite comparable to bonobo and chimpanzees (Addessi et al., 2011). However, although there is suggestive evidence of capuchins being less impulsive and more deliberate than squirrel monkeys, no direct comparison has been made between these two species using the same tasks. One purpose of the study presented in this chapter is to conduct such a direct test.

An ecological reason underlying this potential difference is that, although squirrel and capuchin monkeys are both cebid monkeys, sharing habitats and forming mixed species groups in both their natural environment and our research site, and both have omnivorous diets and large Brain-Body-Ratio, nonetheless capuchin monkeys are extractive foragers and occasional flexible tool-users, whereas squirrel monkeys do not show such diverse manipulating behaviours, requiring delayed access to food, as capuchins (Addessi et al., 2011; Anderson et al., 2010; Fragaszy et al., 2004).

The aim of the current chapter is to explore the potential trade-off between object representation and reaction speed in squirrel and capuchin monkeys. We examined the "Cursory squirrel monkey vs. deliberate capuchin monkey" hypothesis by assessing and comparing the two species' behaviour in terms of reaction speed, accuracy, and precision, in a touchscreen-based target-hitting task. Given the trade-off rationale we mentioned above, this hypothesis, if true, would lead to several predictions. First, the squirrel monkey will reach out for a target significantly faster than the capuchins will. Second, compared to the capuchin monkey, the squirrel monkey will be more likely to interact with non-target objects, or in other word, be less accurate. Third, the capuchin monkeys will reach for the target more precisely, namely the hitting point will be closer to the centre of the target. The three predictions are non-exclusive. Either or both cases being true will support our hypothesis. Note that there must be at least one of these prediction being true if the squirrel monkeys do tend to "act before thinking". Thus, if none of these predictions is true, the "Cursory squirrel monkey vs. deliberate capuchin monkey" hypothesis will be rejected.

Regarding the exploratory nature of the current experiment, the findings are expected not only to explain our previous results, but also to provide a reference for setting up experimental parameters for future studies involving these two species.

3.2 Method

3.2.1 Participants

Although most of the adult monkeys in this research site have had at least some previous experience on touch screen cognitive tests, the apparatus and layout were very different from the current study, especially in terms of the way in which the screen was installed and presented to them. In the previous studies, the screen was placed on a trolley table, so that the monkeys could inspect the environment as usual while watching the screen, and they could only touch the screen when the experimenter allowed them to do so (by pushing the trolley forwards). In the current study, the screen was directly installed on the cubicle window. This might have at least two consequences: (1) it was more difficult for the monkeys to inspect the environment, which might make some individuals feel stressed; (2) the screen was much closer to the subject, and it was always accessible to the subject, so that some monkeys might feel intimidated, whereas some of them might be attracted/distracted by the apparatus itself instead of the stimuli presented. Given these differences, we conducted a familiarisation to screen the subjects. Monkeys who were calm and confident to correctly interact with the stimuli on the screen were included in the test phase.

Six (out of ten) squirrel monkeys and four (out of eight) capuchin monkeys passed the familiarization and participated in the experiment. Four squirrel monkeys and two capuchin monkeys quitted (asked to leave in the middle of sessions, e.g., staying in the corner, knocking or trying to open the cubicle door, etc) mainly because they lost interest very soon or became frustrated/stressed during the task. Another two capuchin monkeys seemed to be involved in a series of fighting, so that other individuals distracted them from the task or even prevented them to entering the cubicle.

For the test phase, apart from one squirrel monkey who stopped coming without specific reason after completing the familiarization, five squirrel monkeys and all four capuchin monkeys completed all ten sessions. All of these monkeys had had previous experience on touch screen cognitive tests.

3.2.2 Apparatus and materials

The experiment was run on two seventeen-inch Elotouch touchscreens (Resolution: 1024 width x 1280 height), one for each social group of monkeys (one for the east monkeys and one for the west ones). Each of the touchscreens was embedded in a cubicle window frame in portrait orientation, so that a subject was facing the screen when it came into the research cubicle. During the experiment, the touchscreen was connected to a laptop computer from which the experimenter controlled the program and presented stimuli on the touchscreen. Two stimuli were used in the experiment: a green grass patch (radius = 80 pixels) was used as the fixation, and a cricket picture (80 x 80 pixel^2) was used as the target. Both stimuli were presented on a rounded “response field” (radius = 472 pixels, centre at [512,512] from the bottom-left) on the screen. All the areas outside this field were grey and unreactive. A Gopro

Hero 4 camera was attached to the cubicle ceiling to monitor and film the monkey and the screen.

Rewards were raisins cut into small pieces (approximately 3mm x 3 mm x 3mm, a half or a third of a whole raisin). Every time a subject succeeded a trial, a reward was given to it through the middle of the gap beneath the touchscreen, namely the gap between the front cubicle window and the cubicle floor.

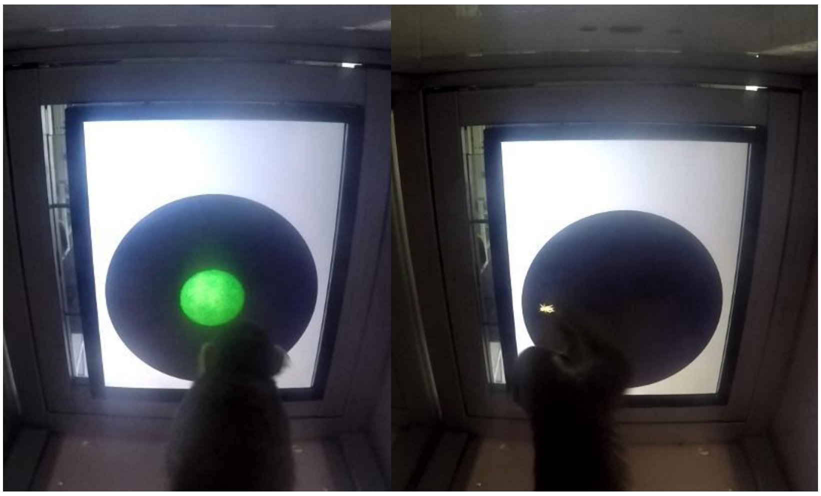


Figure 3-1 Illustration pictures of squirrel monkey participating the task. Left panel: the fixation phase in which the monkey needed to press the green disk (grass patch) to start a trial. Right panel: the "catching" phase in which the green disk disappears to reveal a cricket (on the left of the screen in the picture) that moved towards the edge of the field for the monkey to catch.

3.2.3 Procedure

Familiarization phase

To ensure that a subject was calm, comfortable, and confident to interact with the stimuli (especially the cricket picture) on the touchscreen, we conducted a short training phase comprising ten trials. The training task began with a cricket image jiggling in the centre of the response field. Pressing this "fixation" cricket started a trial: when the centre fixation cricket disappeared with a quick buzzling sound (about 0.5 second) when touched, another cricket image was presented peripherally, as if the cricket fled. The cricket might "reappear" randomly at one of the sixteen locations equally divided on a virtual circle (radius =362 pixels) and jiggle. If a monkey “caught” the cricket (touching the area of 100 x 100 pixel^2 around the centre of the cricket image) within ten seconds, a higher pitch buzzing sound was played and the monkey would be rewarded for this successful “catch”. The criterion of this task was eight successful catches within a ten-trial session. Each monkey had three chances, namely three sessions, to try to pass the training phase and proceed to the test phase.

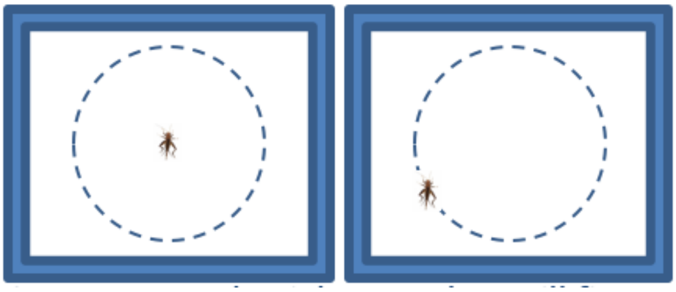


Figure 3-3 Schematic diagram of the training stimuli and procedure on the touch screen. A monkey will first see a fixation cricket in the centre of the screen which requires touching to start the trial (left panel). Then cricket pictures will appear in a random order sequence at one of the locations on a circle (right panel). Note that the dash line is only fictional, and that the size and position are changed for convenience.

Test phase

A monkey who passed the familiarization phase could participate in the test phase on a voluntary basis. Each monkey was expected to take ten sessions, each of which contained fifteen trials.

In each trial, when the monkey was facing and watching the screen the experimenter pressed the “space” button on the laptop to present the monkey with a fixation stimulus, namely, a green grass patch appeared on the middle of the response field, serving as a fixation point. Once the monkey touched the fixation point and lifted its palm, the grass disappeared to reveal a cricket with a quick buzzling sound (about 0.5 second) as if it was previously covered by the grass patch. The cricket then moved outbound towards the edge of the response field, at a speed of 160 pixels/second, directed at one of the sixteen directions equally divided on a circle (see figure 3-2). The monkey was expected to “catch” the cricket before it “escaped”. A successful catch was defined as touching the area of 100 x 100 pixel^2 around the centre of the cricket picture before it disappeared. If a monkey succeeded, a higher pitch buzzing sound was played whilst the cricket disappeared. A piece of reward (a quarter of a raisin) was then immediately given to the monkey by the experimenter. The latency and the precision of each touch after the cricket appeared and before it was caught were recorded along with the direction of the cricket.

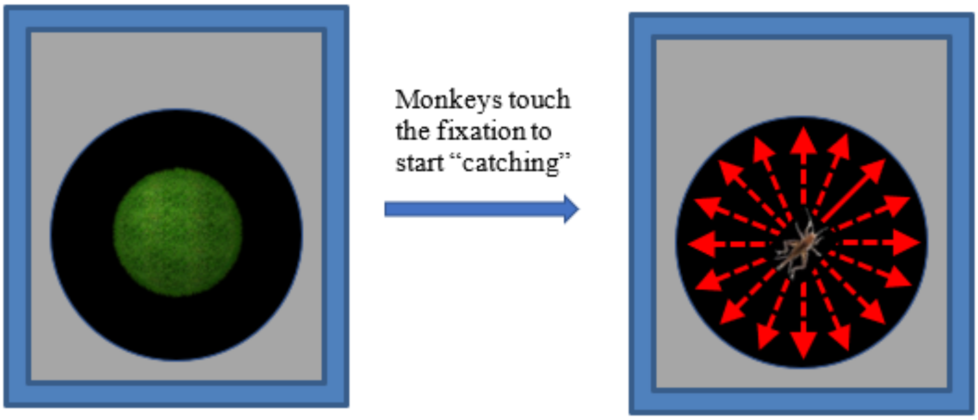


Figure 3-3 Schematic diagram of the stimuli and procedure on the touch screen. The blue rectangles represent the touchscreen. The black part of the screen is the response field, whereas the grey part serves as mere background. The left screen illustrates the fixation phase: the subject needs to touch the green fixation disc to start this trial. After touching the fixation, the "catching" phase started (the right screen): the green disc disappears to reveal a cricket beneath it, moving towards one of the sixteen candidate directions (right arrows) until it disappears into the grey area of the screen. Note the stimuli (fixation and the cricket) are enlarged for convenience.

3.2.4 Analysis

We considered the Species and the Practice effect as two factors of the model. Besides, three dependent variables were considered: (1) *Reaction time* (RT) was defined as the latency between the cricket appearance and the successful "catch"; (2) *Deviation* or distance (pixels) between the hitting point and the centre of the target was calculated to represent how precisely the subject hit the target; (3) *Accuracy* was defined as number of the off-target hits before a successful hit was coded in each trial (lower scores mean greater accuracy).

3.3 Results

Reaction time

The GLMM revealed a significant main effect of Species (F(1,6) = 6.994, p = 0.033, η_p^2 = 0.500). Capuchin monkeys in general “caught” the “cricket” faster. Meanwhile, there was a significant Practice effect (F(1,7) = 2.461, p = 0.018, η_p^2 = 0.260) but non-significant Interaction (F(1,7) = 0.337, p = 0.959, η_p^2 = 0.046). Both monkeys improved their performance to catch the cricket faster over time and their improvement seems to be comparable (see figure 3-4).

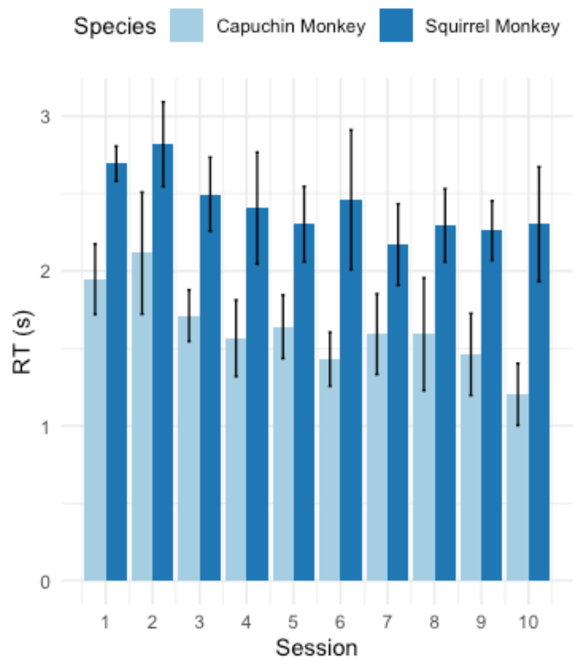


Figure 3-4 Monkeys ’mean reaction time (seconds) to “catch” the target “cricket”

Precision

The GLMM revealed a significant main effect of Species (F(1,7) = 7.23, p = 0.031, η_p^2 = 0.510). The capuchin monkeys hit the target more precisely than the squirrel monkeys. Again, there was a significant Practice effect (F(1,7) = 2.534, p = 0.015, η_p^2 = 0.266) but no significant interaction (F(1,7) = 0.782, p = 0.634, η_p^2 = 0.100). Both species of monkeys became more precise over sessions and we found no difference between their improvement (see figure 3-5).

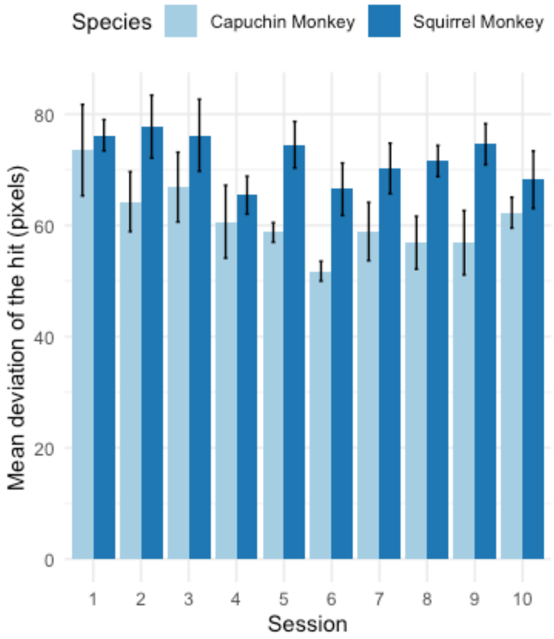


Figure 3-5 The mean distance (pixels) from the monkeys' hit point and the centre of the target "cricket"

Accuracy

Neither main effects nor interaction were significant (Interaction: F(1,7) = 0.686, p = 0.719, η_p^2 = 0.089; Practice: F(1,7) = 0.073, p = 0.680, η_p^2 = 0.094; Species: F(1,7) = 1.031, p = 0.344, η_p^2 = 0.128). We didn't find significant differences between species. In general, most of the monkeys, no matter which species, tended to make averagely two or three off-target hits before they successfully caught the "cricket", whilst their performance didn't improve significantly through ten sessions of practice (see figure 3-6).

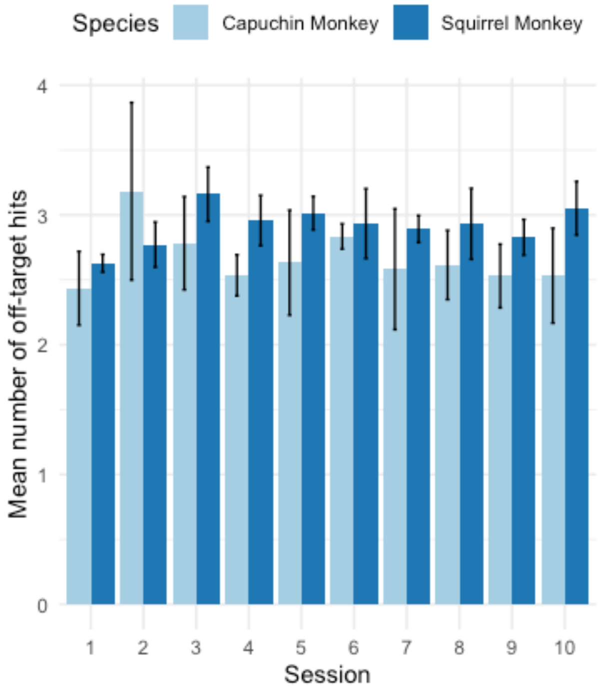


Figure 3-6 The mean number of hits the monkeys missed before they catch the target “cricket”

3.4 Discussion

The “Cursory squirrel monkey vs. deliberate capuchin monkey” hypothesis predicted shorter reaction time, more off-target hits, or/and larger deviation between the hit and the target in squirrel monkeys than in capuchin monkeys. However, the findings didn’t support this hypothesis, and even seemed to go in the opposite direction in some respects. Critically, the squirrel monkeys were significantly slower than the capuchins. In addition, there was no evidence that the squirrel monkeys made more off-target hits than the capuchins. The precision of hits showed that the two species reached the target comparably well, despite their body size difference.

It is notable that the capuchin monkeys can be about 30-40% larger than the squirrel monkey in terms of their body length or height. The larger body size may facilitate their performance on such task, e.g., the capuchin monkeys may benefit from their body size to catch the cricket faster, more accurately, and more precisely. This may involve a widely existing “dilemma” in comparative psychology. On the one hand, a researcher can use the identical display and apparatus to test two species, risking a potential critic that the performance difference is due to some morphological discrepancies (i.e., body size in this context) rather than the difference in cognition. On the other hand, researchers can modify their apparatus and display accordingly at the expense of compatibility. In the current experiment, we choose the first option in order to explain our squirrel monkeys' data in chapter 2 in comparison to the previous study using identical settings on capuchin monkeys (Kersken et al., 2020). Nonetheless, we would like to suggest that it would also be a meaningful supplement for future research to scale the stimuli and display to fit the body size of each species, when comparing these two species in terms of their reaction pattern/tendency.

It could be that squirrel monkeys behave more impulsively or quickly when reaching for real food instead of computerized pictures on the screen. Hasty foraging may have an advantage in the natural environment of monkeys. It is sometimes more important to secure the food before recognizing the quantity or the kind of food, especially in some scenarios such as hunting insects. Thus, we used an insect picture as the target stimulus in the current experiment to increase the ecological validity. Yet, it is possible that the monkeys did not perceive the task as the same as real foraging, given that they receive food rewards from the experimenter instead of actively foraging themselves. The monkeys received real food items after they correctly hit the target cricket, whereas in natural environment or in experiment of chapter 2, they were interacting with or were anticipating real food items. This difference may lead to a lower level of motivation to engage in the current task. Besides, anticipating the delivery of the food rewards may distract the monkeys' attention from the task, and therefore increase the executive demands of the task. For instance, we observed that sometimes a monkey will look towards the seam of the cubicle window where the food would be delivered during the task. And this may be more difficult for squirrel monkeys than for capuchin monkeys given their performance on delay of gratification (Addessi et al., 2013, 2011; Anderson et al., 2010; Evans et al., 2012).

In sum, the capuchin monkeys seem to outperform squirrel monkeys in this task. They reacted as precisely and accurately as squirrel monkeys, but significantly faster. This pattern provides no support for the "cursory squirrel monkey" hypothesis. According to this hypothesis, the squirrel monkeys would tend to launch actions *in parallel with* tracking and representing objects, and this could explain why they failed the "magic box" tasks because they reacted too hastily. However, the direct comparison conducted in the current chapter with a simpler "touch the target" task shows that the squirrel monkeys "caught" the target crickets slower than the capuchins. Moreover, if the hypothesis that squirrel monkeys are hastier and more cursory was true, more erratic touch(es) would have been expected. No such difference was found in the current experiment. The number of off-target hits and the distance between the hitting points and the target centre are both comparable across species. Nonetheless, note that the rewards/targets were visible to the subjects since revealed. It may be arguable that the object representation may not be necessary in this task. Thus, we suggest that these results should be cautiously interpreted within the scope of action capability and behavioural pattern/tendency. In addition, it would be an intriguing future direction to further explore the monkeys' performance in scenarios that the rewards moved behind occlusions.

In the next chapter I will consider two further explanations for the difference between capuchin and squirrel monkey performance on the magic box task. First, a simple explanation is that the capuchin monkeys are better at object cognition in general, and in particular better at individuating objects than squirrel monkeys. Namely the squirrel monkeys' failure in chapter 2 reflects a true object individuation incompetence. Alternatively, it is still possible that both the "Magic box" and the current "Whack-a-cricket" tasks have put squirrel monkeys at a disadvantage because of the particular apparatus and/or task design used may have impeded the squirrel monkeys from presenting their abilities. For example, it is possible that the capuchins benefited from their bigger body size. The capuchins, which are bigger than the squirrel monkeys, might have found it easier to reach for the target cricket. The apparatus, being of the same size for both, might also have required the squirrel monkeys to distribute their attention across a space relatively larger for them. Similar issues could exist in the "Magic box" experiment. The same apparatus and experimental set-up, on the one hand, enabled us to directly compare the performance of the two species, but on the other hand, might bring in extra task demands for squirrel monkeys. Therefore, to assess the squirrel monkeys' competence on individuating objects, we further designed an experiment in Chapter 4, to evaluate the ability of squirrel monkeys to individuate objects, which might reduce some of the demands associated with tracking a reward over a large, occluded distance posed by the magic box task.

Chapter 4 Squirrel monkeys' object individuation revisited: Two-barrier task

4.1 Introduction

In chapter two, we found that squirrel monkeys failed those object individuation tasks which twelve-month-old human infants and their close relatives, capuchin monkeys, passed (Kersken et al., 2020; Van De Walle et al., 2000; Xu & Baker, 2005). We suggested two alternative explanations apart from incompetence. One possibility is that the squirrel monkeys are better at "acting" than "thinking". Namely they prefer to react, visually or manually, as soon as they can, and maybe redundantly, before intaking and processing the information available. In chapter 3 we explored the first possibility, finding little support for it, as squirrel monkeys were not faster and less accurate than capuchin monkeys in a simple target-hitting task. The second possibility is that the "magic box" task was too complex for the squirrel monkeys. The box is relatively large in relation to their smaller body size compared to capuchin monkeys. The task demands may be increased given the nonlinear trajectories of the objects across the scenario, including movements from the base to the top of the box and then the inferred dropping trajectory back to the opening at the bottom of the box. And there was only a single location, or region of interest, for the monkeys to search, which might have provoked a ceiling effect in searching and looking, masking any individuation skills.

To further investigate squirrel monkeys' ability to individuate objects by spatiotemporal or property/kind information, we designed an additional follow-up experiment, with a simpler version of the object individuation task, adapted from Xu and Carey's developmental study (1996). In this new task, objects were moved horizontally on linear trajectories around and/or hidden behind two barriers. There are at least three reasons why the new tasks would be more suitable and less demanding to test squirrel monkeys.

First, we used barriers instead of boxes. It has been reported that squirrel monkeys are more willing to interact with barriers or screens than boxes (de Blois et al., 1998). A possible explanation is that containment events can be more difficult to deal with than occlusion events, even when the variable to represent and reason about is the same (the so-called “Décalages” across event categories, Baillargeon et al., 2012). For example, in two similar object permanence tasks in which “height” is the critical variable to determine whether a target object should be seen, human infants succeed in the occlusion task by 3.5 months while they didn’t pass the containment version of task until 7.5 months (Baillargeon & DeVos, 1991; Hespos & Baillargeon, 2001a, 2001b). Besides, the size of the “magic box” may be too big for the squirrel monkeys, demanding a bigger effort in terms of shifts in spatial attention to the successive relevant events at the top and the bottom of the box. Therefore, we expected the use of spatially closer barriers instead of box(es) in the present experiments would facilitate the performance of the squirrel monkeys in relation to the target objects, and would be more likely to reveal their competence in object individuation.

Second, the objects were moving in a smaller area following simpler trajectories. In developmental studies, simplified object trajectories decrease the difficulty of object individuation tasks, enabling younger infants to demonstrate their competence with easier versions of the task. For example, some researchers suggest that when using simplified object individuation tasks, younger infants show the ability much earlier than they can pass Xu and Carey’s classic tasks (Wilcox & Baillargeon, 1998a; Wilcox & Chapa, 2004; Wilcox & Schweinle, 2002). Critically, whereas the classic tasks involve an object appearing several times from behind a barrier with its trajectory reversed multiple times, the simplified versions include less appearance, reappearances, and reversions, or even just a single one-way trajectory. With such simpler trajectories, 5-month-olds were reported to succeed in property/kind individuation tasks (Wilcox & Baillargeon, 1998a; Wilcox & Schweinle, 2002).

Hence, we aimed to simplify the trajectory of object movement in our present study to better reveal squirrel monkeys’ potential competence by decreasing the task demands. The food items were moved horizontally along the barriers over just about 35 cm, instead of being lifted from the platform, moved towards the top of the box, and dropped into the top hole of the “magic box” (40 cm high, about the body length of an adult squirrel monkey). Besides, though we didn’t go as far as using a single trajectory as some of the studies mentioned above, there was no more than a single circular movement in each trial, namely a location would not be visited by the object(s) more than twice in a trial.

Third, instead of a single search location as in the magic box, the present tasks involved two search locations. More specifically, instead of a single location with either one or two objects, there were two locations behind which 0 or 1 objects could be hidden. In this classic spatiotemporal object individuation task, object(s) are moved along a continuous or discontinuous path (Spelke et al., 1995; Xu & Carey, 1996). Subjects needed an understanding of the continuity principle to use such information to individuate objects. In contrast, in the box paradigm, the number of objects was given more directly. The objects were presented simultaneously or sequentially so that the difference between one- and two-object conditions was more apparent than in the Spelke et al.'s paradigm that used by Xu and Carey in infants (Spelke et al., 1995; Xu & Carey, 1996).

4.2 Spatiotemporal Individuation

We began with spatiotemporal individuation, which is likely to be a more fundamental skill based on spatiotemporal principles, according to Xu and Carey's hypothesis (Carey & Xu, 2001; Xu & Carey, 1996). As an example, in human development, it is an earlier developing skill to individual objects by spatiotemporal information, or representing a "Spelke object" (Xu & Carey, 1996).

4.2.1 Method

Subjects

Twelve female squirrel monkeys (Mean_{age}=12.75 years; age ranged from 9 to 19 years) participated in the familiarization phase of the experiment. Four more individuals took part in the test phase, making a total of sixteen female squirrel monkeys (Mean_{age}=12.50 years; age ranged from 8 to 19 years). Each subject came into the research cubicle to be isolated for the experiment on a voluntary basis to win supplemental food rewards (sunflower seeds and raisins), without any food or water deprivation. All monkeys could leave and would be let out once they showed any signs of anxiety, frustration, or losing interest, etc. The study was granted ethical approval by the University of St Andrews' Animal Welfare and Ethics Committee.

Apparatus

The apparatus consists of two semi-cylinder barriers placed on a box with a long slit on the top and a wide opening at the rear. The barriers can be knocked down and turn into tunnels, allowing the monkeys to retrieve food in this way (Figure 4-1). The food item(s) were manipulated by the experimenter (i.e., popping up, moving along the slit, resting behind the barrier(s) when the sticks pulled away), via cocktail stick(s) through the slit, whilst keeping his hands invisible in the box to eliminate potential distraction.

This apparatus was placed on a trolley table, so that the experimenter could demonstrate the object-moving events to the monkeys at a distance out of their reach (about 40 cm), and then push the apparatus towards the research cubicle for the monkeys to make choices (see Figure 4-1 and Figure 4-2).

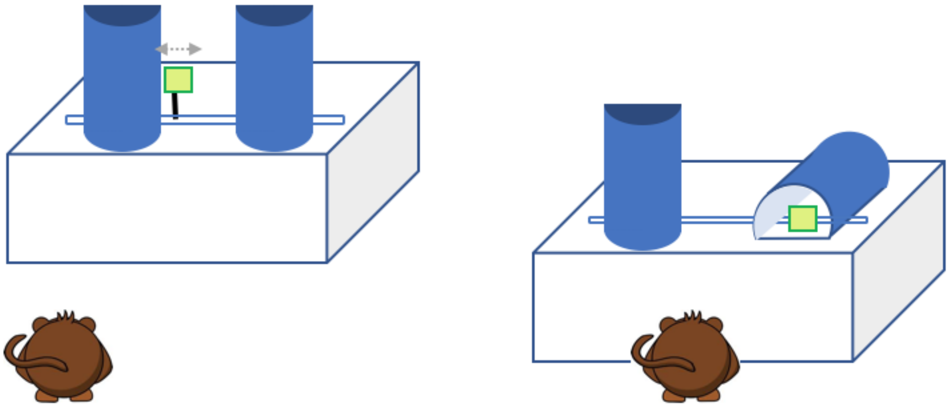


Figure 4-1 Illustration of the set-ups. Left: Moving the food item(s) along the slit and then hiding it when both screens stand. Right: Pushing the apparatus towards the monkey to allow search – the monkey can knock over the screen(s) to reveal and access the hidden food.

Procedure

The task involved monkeys searching for food in one of two locations after watching the experimenter moving the food item(s) along the box surface until it came to rest behind the two barriers. All squirrel monkeys were free to participate in the study, that took place in the research cubicles. A monkey received one sunflower seed after entering the cubicle for participating. It then watched the experimenter's demonstration moving and hiding raisin(s), before it searched for these higher value food items. In each trial, there were either one or two raisins, whilst the path and order that the objects moved around varied across two familiarization conditions (see Figure 4-2) and four test conditions (see Table 4-1).

Familiarization phase

This phase was to ensure the monkeys were willing and able to interact with the apparatus, specifically, knocking down the barriers to retrieve food behind. The monkey saw one piece of food placed behind one of the screens in the *Single-object* condition, whereas in the *Two-object* conditions it may see two pieces of food simultaneously presented and then placed, one on each side, one after another. In both conditions, the experimenter showed the food item(s) on the stick(s) at the edge(s) of the stage, drawing the monkey's attention, before moving them behind the barriers with sticks (see Figure 4-2). The stick(s) was then pulled out to leave the food on stage before the experimenter pushed the whole apparatus towards the monkey. Thus, the monkeys couldn't access the sticks at any time during the experiment. Afterwards, the monkeys could search for up to twenty seconds or until they searched both locations. The experimenter helped the monkeys to get the food if the choice was correct when necessary. Any missing food items would be removed after the trial.

Each monkey received up to ten trials per session for two mixed sessions, thus completing twenty trials including ten One-object and ten Two-object trials. The hiding locations and order were counterbalanced so that each barrier was equally highlighted.

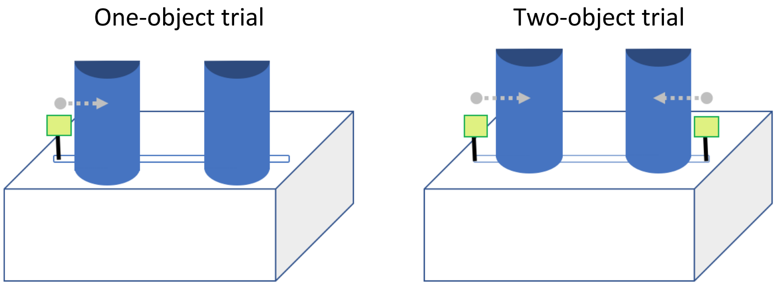


Figure 4-2 Illustration of the familiarisation conditions. Left: One object condition. Food will be placed in the left (illustrated here) or right. Right: Two objects condition. Two locations will be baited sequentially (left-first illustrated here) after both food items are shown to the monkey.

Note that there was only one correct location in a One-object trial, which at the same time was also the last highlighted location, while in Two-object trials both locations were "correct" as both had food. Therefore, if the monkeys simply relied on some attentional rule to search the last highlighted location, we should find similar location preferences in both conditions. Otherwise, if the monkeys did represent up to two objects based on object permanence, there should be only weak or even no preference for location in the Two-object trials.

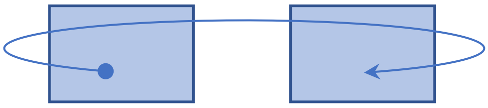
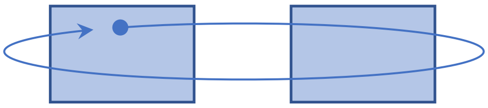
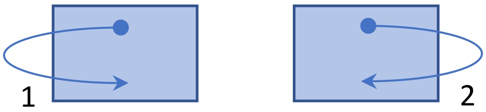
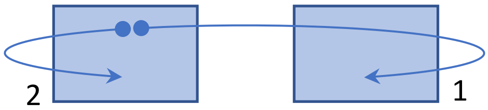
Test phase

The general procedure of each trial is similar to that in the familiarization phase, except for the demonstration and preparation. Instead of a simple hiding event, the experimenter moved the raisin(s) along certain different paths according to the conditions. Specifically, in a *Continuous* trial, it was one raisin moving along a continuous path, whereas in a *Discontinuous* trial, it was two raisins moving successively but along a discontinuous path. After each trial, the experimenter reset the apparatus and placed raisins for the next trial in their "start" points behind the barriers, with his back completely blocking the monkeys' sight. Note that for each condition, the movement speed of the food item(s) was the same. Moreover, for each pair of conditions (Continuous A vs. Discontinuous A; Continuous B vs. Discontinuous B), the duration of the movement was the same: in the discontinuous trials the experimenter waited before the gap, as if the object took this period of time to move from one barrier to the other.

Monkeys received forty trials in two *Continuous* conditions and two *Discontinuous* conditions (10 trials per condition), mixed and distributed in five eight-trial-sessions with counterbalanced order. The side of the "correct" location(s) with food, the last highlighted location, and the starting direction of the movement were balanced between *Continuous* and *Discontinuous* conditions. The critical differences between conditions are illustrated in Table 4-1 with an example of each condition.

Similar to the familiarization phase, we predicted that a competent monkey would first search the last-visited location, which was also the correct location, in *Continuous* trials, whereas in *Discontinuous* trials, they would have only weak or even no location preference if they represented one food item behind each barrier.

Table 4-1 The objects' moving path and expected performance of the experimental conditions.

Condition	Moving path of the food item (s)	Object Number	Plausible choice	Last location	Starting Direction
Cont. A		1	Right	Right	Left/outwards
Cont. B		1	Left	Left	Right/inwards
DisC. A		2	Both	Right	Left/outwards
DisC. B		2	Both	Left	Right/inwards

Note: Cont. = continuous; DisC. = discontinuous. The number in the diagram indicates the order of moving events (e.g., for DisC. A, the first food item comes out from the left barrier, goes back, and then the second food item appears and disappears on the right side). All movements involved in the test were linear rather than circular as illustrated above.

The food item starts from either the left (as illustrated above) or the right with counterbalanced order. In addition, we counterbalanced several key features, including the number of food items, the correct location(s) for the food, the last visited location, and the starting direction of the movement.

Moreover, by using two types of moving paths under each continuity condition, the present study aimed to get a fuller picture of how squirrel monkeys understand spatiotemporal rules. We designed a critical condition, Discontinuous B, in which a fully competent monkey should have no or only a weak preference between the two locations. Thus, with the understanding of the continuity principle, an individual may perform differently across Discontinuous B and the continuous conditions, in which we anticipate the monkeys to search the last-visited location first as this is always a safe choice and it is where the monkeys most likely attend to. More specifically, including the Discontinuous B condition may help us to rule out two alternative possibilities. One is that an individual can establish an association between the movement of the target between the two barriers and the fact that only the last-visited-location contains food. The Discontinuous B condition included such feature as well as the Continuous A and B trials (see Table 4-1). In summary, a significant difference between the performance on Discontinuous B and that on the two continuous conditions would support the conclusion that the monkeys use the continuity principle (rather than an associative rule) when individuating objects.

Another possibility is that an individual can rely on the overall routes that object(s) ever moved along during the whole demonstration, while ignoring the temporal order. Specifically, the overall trace of object movement is very similar between Continuous A and Discontinuous B in terms of the purely spatial information (see Table 4-1). Only the temporal continuity distinguishes the two conditions. Thus, an individual who takes the temporal information into account will notice that the objects actually moved discontinuously, and thus there should be two objects. In sum, we argued that a species that is equipped with a full-fledged understanding of spatiotemporal principles should react significantly differently across Continuous and Discontinuous conditions regardless of the movement types, i.e., we expected a main effect of Continuity but no interaction in the ANOVA analysis.

Measures and coding

We used three dependent variables to measure the monkeys' performance in the test phase. First of all, we coded where or which location a subject searched first, or the *First Search*. Specifically, we focused on the last-visited location (namely the location that the object moved into at last), which in any case contained food, namely, a "safe choice". In continuous trials (or 1-object trials), there was probably only one piece of food and it should be hidden behind the last-visited-location. However, there could be a second piece of food hidden behind the other location in the discontinuous trials (or 2-object trials). Thus, if a monkey was capable of individuation, it would be more likely to search the last-visited location first in the continuous trials (or 1-object trials) compared to the discontinuous trials (or 2-object trials). In the latter case, we predicted that a monkey might start searching with either location because both locations might contain food.

In addition, we also coded whether a monkey would search or not the other location after they retrieved the food item from the last-visited-location. According to the rationale of previous studies using the box paradigm (e.g., Mendes et al., 2008), an animal would search after an initial retrieval of food if they thought that there was more than one food item in this scenario. Therefore, if a squirrel monkey in our experiment represented two objects, it should search both locations. More specifically, in this case, it is likely that the monkey searched the last-visited-location (the "safe choice") first and then searched the other location. In contrast, we expected to find less second search behaviour in the continuous trials (or 1-object trials). A fully competent monkey would probably stop searching in this kind of context after it found a piece of food in the last-visited-location. Note that this measure can be used as a more robust evidence that an animal individuates two objects on the one hand, while on the other hand this may be more stringent and demanding compared to the *First Search* measure we mentioned above. For example, it may be easier for a monkey to show a preference between two locations by searching immediately ("*First Search*") after watching the event, whereas it may cost this monkey more cognitive resources to remember that there were two objects and to maintain this representation until it retrieves the first object, consumes it, and gets ready for the second search.

Besides, we also considered the percentage of trials that a monkey searched both locations regardless of which one was chosen first, or the "*Thorough Search*". We include this analysis in analogy to the Magic Box experiment: when there is only one location, searching after retrieval may be explained by either searching for the missing object, or searching because of a higher level of arousal. Hereby in the present two-barrier task, we can distinguish these two behaviours with separate measures: searching for the missing object leads to more *Second Search* behaviours (but not *Thorough Search* behaviours), while searching because of arousal leads to more *Thorough Search* behaviours. Note that the difference between the *Thorough Search* and the *Second Search*, is that the "*Thorough Search*" is a general measure of to what extent a monkey would like to search both locations, whereas the "*Second Search*" is a conditional measure of whether a monkey searches again *after* retrieval.

4.2.2 Results

Familiarization phase

Twelve monkeys participated in the familiarisation phase, seven of which completed all twenty trials whereas the rest left early because of distraction or losing interest.

Table 4-2 monkeys' performance in the familiarisation phase

Subject	Age	Number of trials			Searching 1st location		Searching 2nd location	
					1-object	2-object	1-object	2-object
		1-O	2-O	total	Correct	Last-visited	2nd search	2nd search
Toomi	14	8	6	14	0.88	0.33	1.00	1.00
Gab	9	10	10	20	0.60	0.50	0.67	1.00
Cali	15	10	10	20	0.90	0.40	0.22	0.60
Roca	17	10	10	20	0.70	0.50	0.43	0.80
Maya	17	10	10	20	0.60	0.40	0.83	0.90
Tatu	19	6	4	10	0.83	0.50	0.20	0.25
Pelusa	9	10	10	20	0.90	0.40	0.67	0.80
Elie	15	6	4	10	0.83	0.50	0.80	0.25
Ciara	9	5	6	11	0.80	0.50	0.75	0.83
Amar	9	10	10	20	0.70	0.50	1.00	1.00
Dora	10	10	10	20	0.70	0.60	0.71	1.00
Lexi	10	4	8	12	0.50	0.88	1.00	0.63
Mean	*12.75*	-	-	-	*0.75*	*0.50*	*0.69*	*0.76*

Note: 1-O=one-object condition; 2-O=two-object condition. The proportion of 2nd search = the number of trials a monkey searches the second location after they successfully retrieved the first food divided by the number of completed trials in which the monkey successfully got the first piece of food at her first attempt.

In the *One-object* trials, on average in seventy-five percent of the trials, the monkeys first searched the last-visited location and found the food. In contrast, monkeys did this only in fifty percent of the Two-object trials, showing a significantly weaker preference as we expected (t(11) = 3.426, p = 0.006, d = 0.989). Interestingly, after a monkey found the first piece of food, a second search in the other location was equally likely to happen, with no significant difference between One- and Two- object trials (One-object: 69.0%; Two-object: 75.5%; t(11) = -0.788, p = 0.447, d = -0.227). It is arguable that such finding may suggest that the monkeys didn't expect two objects in the two-object trials. One possible explanation is that in the familiarization task, both objects were visually accessible to the subject (in the 2-object trials), which was likely to distract the monkey from representing both objects that appeared simultaneously but distantly from the beginning. Hence, it is possible that in some 2-object trials, the subjects didn't represent one of the objects so that their second search behaviours were not significantly more than that in the 1-object trials. Besides, given the nature of the familiarization phase, it is also possible that the subjects were more cautious to interact with the apparatus. Thus some of the subjects might not search again regardless of the condition (i.e., even when they recognized two objects). Moreover, the data may be more noisy so that the difference (supposedly there is indeed a difference regarding the data and the effect size) between conditions may be covered (false negative). Taken together, we refrained to make further discussions based on this result.

Test phase

Sixteen squirrel monkeys participated in the test phase and finished all the trials. Compared to the familiarisation ($N = 12$), four more monkeys joined in later during the test phase (i.e., without familiarization) and completed the task. Thus, all of these sixteen monkeys were included in the repeated ANOVA analyses. Each analysis considered the continuity of the path (Continuous vs. Discontinuous) and the type of movement (A vs. B, see Table 4-1). Besides, when analysing the First search and the Thorough search performance, the trial number was also included in the model as a co-variate in order to control the potential practice effect. We didn't consider the practice effect when analysing the second search, since this variable, aiming to measure whether a subject searches the first location after retrieving food from the last-visited location, is defined by a conditional event. Namely, the second search score is subject to the first search score and it only makes sense when considering the percentage of trials., An analysis based directly on the number of trials or a trial-by-trial analysis may lead to errors. Here is a hypothetical example. In ten trials of condition A, a subject always searched the last-visited location first and searched the other location afterwards. Meanwhile, in ten trials of condition B, this subject did so only in five trials, did the other way around in two trials, and searched only the last-visited place in three trials. If we use the frequency of trials, the ratio of second search trials across conditions is 10 versus 5. However, if we use the percentage of trials and consider the nature of this variable, the ratio should be 10/10 versus 5/8. The problem is that the former score misses the precondition of searching the last-visited location first. Therefore, we only considered the practice effect on a trial-by-trial basis in *First search* and *Thorough search* analyses, but not in *Second search* analysis.

First search

A 2-by-2 repeated measure ANOVA was used with continuity of the path and the type of movement as the two factors. The monkeys first searched the last-visited location significantly more often in the Continuous trials than in the Discontinuous trials, $F(1, 15) = 33.601$, $p < 0.001$, $\eta_p^2 = 0.691$ (Figure 4-3). There was no significant main effect of movement types ($F(1, 15) = 3.571$, $p = 0.078$, $\eta_p^2 = 0.192$) nor significant interaction ($F(1, 15) = 3.871$, $p = 0.068$, $\eta_p^2 = 0.205$). Besides, none of the main effect or interactions involving Practice was significant (Figure 4-4; Practice: $F(1, 9) = 0.758$, $p = 0.655$, $\eta_p^2 = 0.048$; Practice * Continuity: $F(1, 9) = 0.818$, $p = 0.052$, $\eta_p^2 = 0.052$; Practice * Type: $F(1, 9) = 1.791$, $p = 0.075$, $\eta_p^2 = 0.107$; Practice * Continuity * Type: $F(1, 9) = 0.696$, $p = 0.712$, $\eta_p^2 = 0.044$).

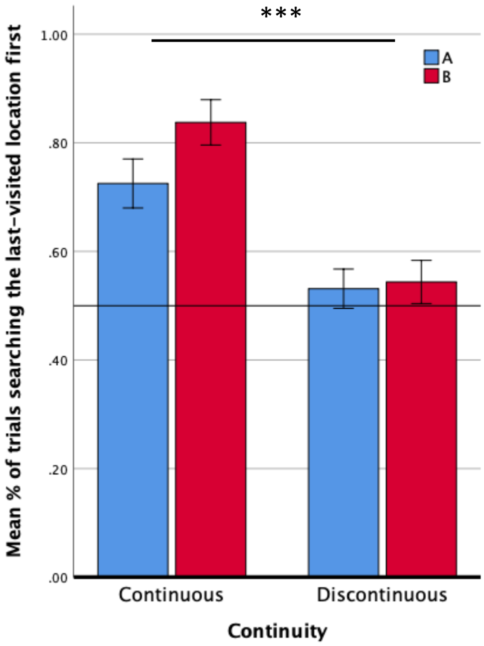


Figure 4-3 The squirrel monkeys' "first search" performance across conditions. The "first search" measures which location a monkey searched in a trial. A and B represent two different pairs of movements.

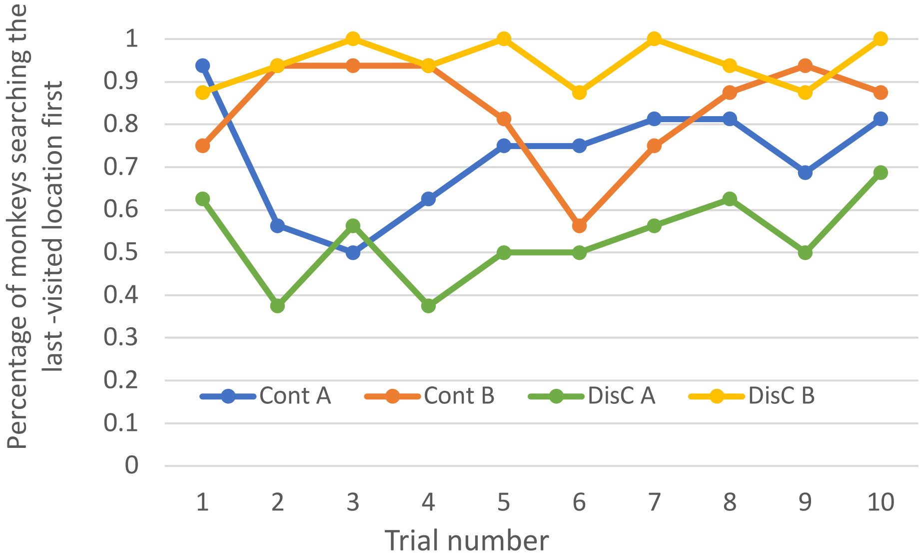


Figure 4-4 The squirrel monkeys' "first search" performance across trials.

Second search

We also coded the "second search", defined as searching another location after successfully retrieving food from the last visited location, which is in line with the rationale of previous studies using the box paradigm (Mendes et al., 2008). A 2-by-2 repeated measure ANOVA revealed a significant main effect on Continuity (F(1, 15) = 6.663, p = 0.021, $\eta_p^2 = 0.308$). In Discontinuous trials, the monkeys were more likely to search the other location after they retrieved the food from the last-visited location first (figure 4-5). There was no significant interaction (F(1, 15) = 0.031, p = 0.862, $\eta_p^2 = 0.002$) nor main effect on Movement Types (F(1, 15) = 0.426, p = 0.524, $\eta_p^2 = 0.028$).

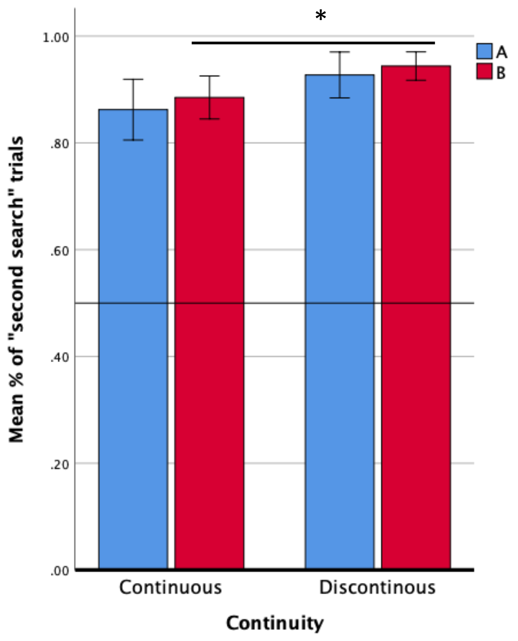


Figure 4-5 The squirrel monkeys' "second search" performance across conditions. The "second search" measures whether or not the monkey continued to search the other location after successfully retrieving the food from the last-visited-location. A and B represent two different pairs of movements.

Thorough search

We also considered the percentage of trials that a monkey searched both locations regardless of which one was chosen first. A 2-by-2 repeated measure ANOVA with continuity of the path and the type of movement as two factors revealed that the monkeys searched marginally more often in the Discontinuous trials, $F(1, 15) = 4.401$, $p = 0.053$, $\eta_p^2 = 0.227$. There was no significant main effect on movement type ($F(1, 15) = 0.143$, $p = 0.711$, $\eta_p^2 = 0.009$) nor interaction ($F(1, 15) = 0.082$, $p = 0.779$, $\eta_p^2 = 0.005$). Besides, none of the main effect or interactions involving Practice was significant (Practice: $F(1, 9) = 1.326$, $p = 0.229$, $\eta_p^2 = 0.081$; Practice * Continuity: $F(1, 9) = 0.936$, $p = 0.496$, $\eta_p^2 = 0.059$; Practice * Type: $F(1, 9) = 0.830$, $p = 0.590$, $\eta_p^2 = 0.052$; Practice * Continuity * Type: $F(1, 9) = 1.212$, $p = 0.293$, $\eta_p^2 = 0.075$).

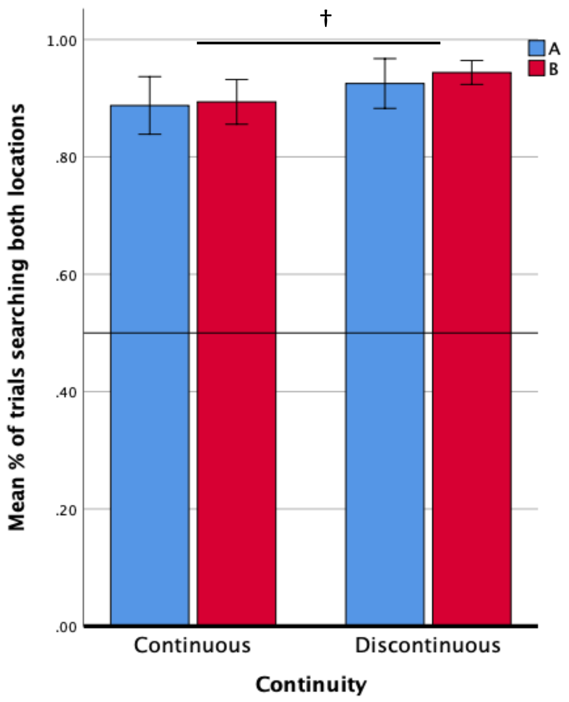


Figure 4-6 The squirrel monkeys' "thorough search" performance across conditions. The thorough search measures whether a monkey searched both two locations in a trial. A and B represent two different pairs of movements.

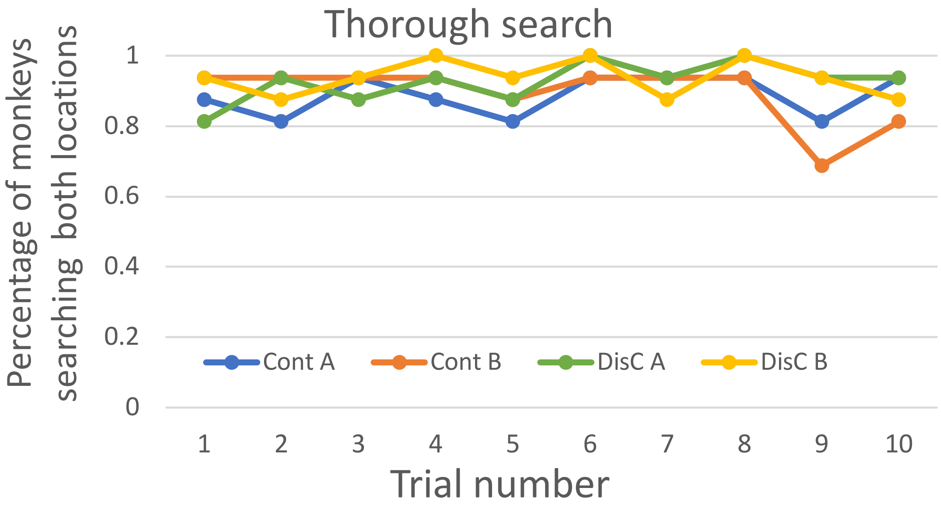


Figure 4-7 The squirrel monkeys' "Thorough search" performance across trials.

4.2.3 Summary

The present experiment adapted a two-location paradigm to investigate whether squirrel monkeys were able to individuate objects based on spatiotemporal information. The critical evidence of competence would be the “first search” performance. If the squirrel monkeys can individuate objects spatiotemporally, they would represent only one object when the moving path was continuous, preferentially searching the last-visited location first, but represent two objects each hiding behind one of the barriers when the path was discontinuous and therefore searched the two locations without preference. In accordance with this rationale, our findings suggested that squirrel monkeys were sensitive to spatiotemporal information when individuating objects, an ability mastered by human infants by the age of 10-month or even earlier (Spelke et al., 1995; Xu & Carey, 1996).

Beside the “first search” measure, we also coded and compared the monkeys’ “second search” and “thorough search” performance across conditions. The results were largely in line with the “first search” measure, confirming an ability for spatiotemporal individuation, though the effect sizes were much smaller. However, it was noteworthy that the squirrel monkeys were very inclined to engage in “second search” or “thorough search” even when there was only one food item involved in that trial (Continuous conditions). In other words, there was likely some “ceiling” effect that undermined the power of the test to reveal differences between conditions. If this is the case, the “first search” performance may be able to serve as a complement to the measures based on frequency or duration of searching, and can reveal the monkeys’ ability more sensitively.

Moreover, the current experiment also included two pairs of movement types to further explore how the monkeys used spatiotemporal information. This helps address potential arguments that squirrel monkeys passed the task based on associative rules instead of the expectation of hidden object(s), or that the representation of objects in squirrel monkeys may be relatively blurry and fragile, given their failure in the magic box experiment presented in chapter 2. In the standard spatiotemporal individuation task (Xu & Carey, 2001), subjects watched only the pair of A type movements (Continuous A versus Discontinuous A). One may succeed in this task with an understanding of the continuity principle, which was the original conclusion of these authors. Yet, one could also rely on an associative rule that only one object can be found when there is a connecting path between the two locations. Or one could simply rely on the spatial continuity of the movements but ignore the temporal discontinuity information. One may also make similar argument(s) from the perspective of attention. Namely, the (dis)continuity itself might draw part of the subjects’ attention to the first-visited location, leading to a more equally distributed choice pattern. Our study included Discontinuous B (and its matched condition Continuous B) in order to rule out these possibilities. Inter-barrier trajectory can be detected in Discontinuous B, as well as in those two continuous conditions. However, such feature alone is not sufficient for successful individuation. An individual who relies on the associative rule or encodes only the spatial aspect of the information will treat the Discontinuous B condition as another one-object condition. And one who passed the task based on an attentional effect caused by spatial discontinuity should perform the same in Discontinuous B compared to Continuous B. We didn't find such evidence in the current experiment (main effect of continuity and no significant interaction between continuity and movement type)[2].

The squirrel monkeys not only encoded the spatial information, but also integrated the temporal information. Thus, our findings revealed that the squirrel monkeys succeeded in the task not by using simple associative rules. Nonetheless, a potential attentional effect is embedded in the present paradigm. The spatiotemporal continuity, especially the temporal continuity, is likely to affect the subjects attention distribution and their behavioural patterns. Such attentional effect may contribute to the object representation process (e.g., noticing the spatiotemporal gap so that one can infer a second object), but it is difficult to fully disentangle them. In the current experiment we matched the display and the features across conditions, and designed two pairs of movement types trying our best to avoid the alternative explanation of attentional effect. Along with the rationale in the previous literature (Spelke et al., 1995; Xu & Carey, 1996), we tend to assume the monkeys' choices were based on expectation rather than attentional effect. Yet, we recognize that the attention interpretation is an important issue for the future studies of this field. Studies using new techniques such as eye-tracking or adapted paradigms including a set of specially designed control conditions (e.g., one or both of the objects leaving the scene after visiting the hiding location) may be helpful to advance our understanding.

Taken together, we conclude that, it is most likely that the negative findings in chapter 2 could be explained by performance problems rather than the lack of spatiotemporal

---

[2] We also conducted a pairwise comparison between Discontinuous B and the two continuous conditions to confirm this. The "first search" results were both significant, surviving the Bonferroni multiple comparison correction.

individuation competence. The smaller and more squirrel-monkey-friendly apparatus that we designed in the current experiment, the linear horizontal trajectories used, and the two-location set-up with choice tasks, may all have contributed to the squirrel monkeys' success in object individuation task.

Next, we designed and conducted a property/kind individuation task with the same paradigm.

4.3 Property/kind individuation

The present study aims to investigate whether squirrel monkeys can individuate objects by using property/kind information. Furthermore, by manipulating the latency time (i.e., the temporal gap between events) we also aim to explore whether this cognitive process will be affected by a spatiotemporal heuristic, the "tunnel effect". Previous studies have shown that when an object A moves into a barrier after which another object emerges from the other side of the barrier in a spatiotemporally continuous trajectory (i.e. without any time delay in the expected movement trajectory of the object), individuals may neglect the featural difference between different pre- and the post-occlusion objects to perceive a single persisting object (Burke, 1952; Flombaum et al., 2004; Flombaum et al., 2012; Scholl & Pylyshyn, 1999; Spelke et al., 1995). Thus, we expected that in our experiment the no-delay condition will be more difficult than the delay condition, i.e., when the second object emerges immediately from the barrier which the first object just moved into, the monkeys would be. more likely to neglect the property/kind difference and take it as a one-object event.

4.3.1 Method

Subjects and Apparatus

The subjects came from the same population at Living Links, Edinburgh Zoo as in the previous study.

Sixteen female squirrel monkeys participated in the familiarization phase of the experiment, and twelve of them took part in the test phase. Each subject came into the research cubicle to be isolated for the experiment on a voluntary basis to win supplemental food rewards (sunflower seeds and raisins), without any food or water deprivation. All monkeys could leave and would be let out once they showed any signs of anxiety, frustration, or losing interest, etc. The study was granted ethical approval by the University of St Andrews' Animal Welfare and Ethics Committee.

The apparatus and materials were the same as in the spatiotemporal task described above, except that two different kinds of food were used, raisins and peanuts. Specifically, to highlight the property/kind difference while matching the size of food items to the best, we used half of a blanched peanut in contrast to a whole raisin in the conditions involving two different items. Besides, food items were manipulated with tiny bamboo clamps (namely, semi-split bamboo sticks, each about 15 cm in length) instead of the cocktail sticks used in the previous task to better control the movement of the peanuts. For convenience, we will still call these tools "sticks", and call half a blanched peanut "a peanut" hereafter.

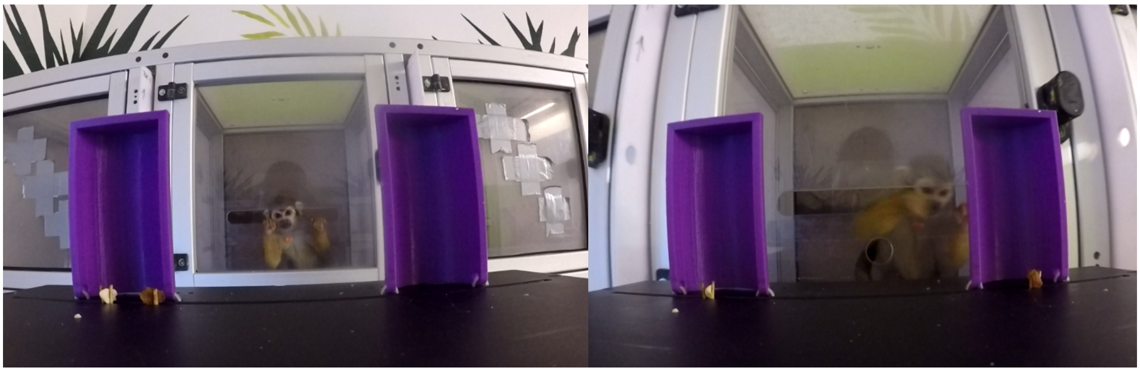


Figure 4-8 The set-up at the beginning of a trial (left panel) and the moment that the monkey began to search (right panel), both from the perspective of the camera. Note that the two food items used in this example: the left is a peanut, and the right is a raisin.

Procedure

The procedure was similar to the spatiotemporal task, except for the demonstration of which food items were used and how they were moved around and hidden behind the barriers.

Familiarization trials

We first conducted a familiarization phase, not only to ensure monkeys were willing and able to interact with the apparatus, but also to set up a baseline of how squirrel monkeys treated different kinds of food rewards. Three conditions were included: single object (Single), two objects of the same kind (Two-Same), and two objects of different kinds (Two-Different). The general procedure was the same as in the spatiotemporal task. A monkey would watch one piece of food moved behind one of the screens in the *Single-object* condition, whereas in the *Two-object* conditions (Same or Different) the monkey would see two pieces of food simultaneously presented and then placed, one on each side, sequentially. The food's location (Side), the kind of food used, or which food item appeared first were counterbalanced in a pseudorandom way.

Each monkey received thirty-two trials, including sixteen Single-object trials (eight with a raisin and eight with a peanut), eight Two-Same trials (four raisin-raisin and four peanut-peanut trials), and eight Two-Different trials (half with a raisin appeared first and half with a peanut appeared first).

We expected the monkeys to search the food's location first in the Single trials, and to search either side equally in the two-object conditions when both sides contained food.

Test Trials

The test trial involved the experimenter using two sticks to move two pieces of food items around the two barriers along the "slit track" cut into the stage. The food items involved in a trial were either the same or different (Sameness: Same vs. Different). In each condition, the movement trajectories of the two objects were similar across trials and could be seen as a spatially continuous path from the monkeys' perspective. However, we manipulated the time lag between the movement of the objects. Namely, the second object appeared immediately after the first one moved into a barrier (Non-delay) or waited for about three seconds before launching. Thus, in half of the trials this temporal gap would make the path discontinuous in time. About half of the monkeys started with the Delay trials, while the other half started with the Non-delay ones. In each block, we counterbalanced and mixed whether the two food items were of the same kind (Sameness: Same vs. Different) and which object appeared the last in each trial (Food type: Peanut vs. Raisin) in a pseudorandom way. Therefore, each monkey received eight conditions of trials in a 2-by-2-by-2 design (Sameness: Same vs. Different; Food type: Peanut vs. Raisin; Temporal gap: Delay vs. Non-delay). Each condition contained ten trials (see Table 4-3 and Table 4-4 for the conditions and the design).

Table 4-3 An example of the mixed blocked design

Subject	Session	Trial1	Trial2	Trial3	Trial4	Trial5	Trial6	Trial7	Trial8
1	1	SAL	SAR	DAR	DAL	SBL	SBR	DBL	DBR
	2	SBR	SBL	SAR	SAL	DBL	DBR	DAR	DAL
	3	DAL	DAR	DBL	DBR	SAR	SAL	SBL	SBR
	4	DBR	DBL	SBL	SBR	DAR	DAL	SAR	SAL
	5	DAR	DAL	SAL	SAR	DBR	DBL	SBR	SBL
2	…								

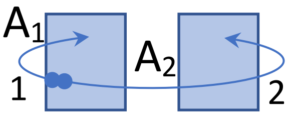
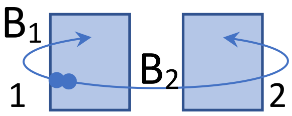
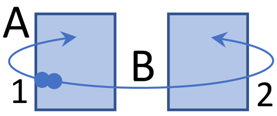
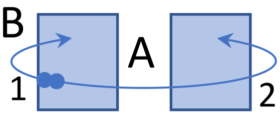
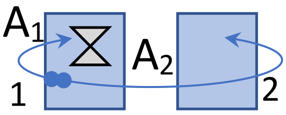
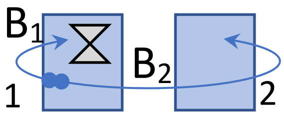
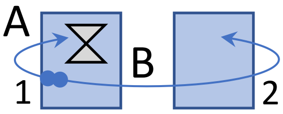
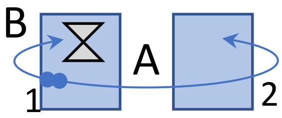
Note: S = same objects; D = different objects; A = moving event ends up with object A; B = moving event ends up with object B; L=movement starting from the left; R= movement starting from the right

In each trial, both objects were surreptitiously placed behind one of the barriers (Barrier A) beforehand. When the trial began, the first object appeared from Barrier A moving OUTWARDS towards the edge of the stage, and then returned back to Barrier A. With or without a delay according to the condition, the second object appeared from Barrier A moving INWARDS towards the other barrier (Barrier B), passed through behind Barrier B without stopping till it reached the other edge of the stage, and finally returned back to Barrier B.

The experimenter then pushed the trolley table forward to let the monkey make its choice. After both barriers were searched or fifteen seconds had elapsed, the trolley was pulled back. The experimenter then came to the middle between the monkey and the apparatus to set up the next trial. If the monkey failed to get a piece of food after searching that location correctly, the experimenter would retrieve it for the monkey. Other leftovers would be removed. The experimenter then hid the food items for the next trial while blocking the monkey's sight with the experimenter's own body.

Similar to the spatiotemporal task, we controlled the duration of each type of movements, depending on whether there was a delay. Namely, the total duration of movement across the four Non-delay conditions was approximately the same. Whereas the total duration of movement in each of the Delay condition was about three seconds longer.

Table 4-4 Illustration of moving patterns (take left-started trials as examples)

Condition		Moving path	Last-visited location	Prediction of 1st search choice
Non-delay	Same A		Right	Right
	Same B			
	Diff. A		Right	Either
	Diff. B			
Delay	Same A		Right	Right
	Same B			
	Diff. A		Right	Either
	Diff. B			

Note: The number in the diagram indicate the order of moving events; A1/A2/B1/B2 refer to different objects; ⧖ *means there is a pause (about 2 seconds) between the first object disappearing and the second object appearing. (For example., for Diff. A with delay, first the food item A comes out from the left barrier, goes back, and then the second food item appears from the same location after about 2 seconds, moves towards and then around the other barrier until it disappears on the right side). The food item starts from either the left (as illustrated above) or the right with counterbalanced order. Besides. we also counterbalanced several key features, including the sameness of the two food items, the food's location, the last visited location.*

The object movements and the subsequent searching of the monkeys were filmed with a Gopro Camera on the rear part of the stage from the experimenter's perspective (i.e., recording the front view of the monkeys). The searching behaviours, especially the location(s) and the order of searching, were recorded and compared between conditions. The critical comparison was "Same" vs. "Different". We expected a competent monkey to search the correct location first in a Same trial, but to search equally between the two locations in the Different trials, as these should be suggestive of there being two pieces of food, one behind each barrier: the one that moved behind the first barrier, and the different one that emerged from that barrier shortly afterwards and moved behind the second barrier. In addition, we anticipated that a tunnel effect could be found, so that this pattern would appear only in trials with a delay, namely, with a temporal discontinuity. These results would imply the squirrel monkeys have a robust object representational ability based on property/kind information.

4.3.2 Results

Familiarization phase

We conducted a one-way ANOVA to compare the three conditions: Single, Two-Same, and Two-Different. Since the number of trials across conditions was uneven, we used the proportion of trials a monkey searched the last-visited location first as the independent variable. The ANOVA revealed a significant main effect of condition ($F(1, 8) = 16.474$, $p = 0.002$, $\eta_p^2 = 0.600$). The post hoc pairwise comparison (LSD) suggested that the monkeys had a stronger preference for the last-visited location when there was only one piece of food (Single - Two-Same = 0.370, $p < 0.001$; Single – Two-Different = 0.245, $p = 0.002$; Two-Same - Two-Different = -0.125, $p = 0.074$). Thus, the squirrel monkeys did differentiate the trials with two objects from those with only a single object.

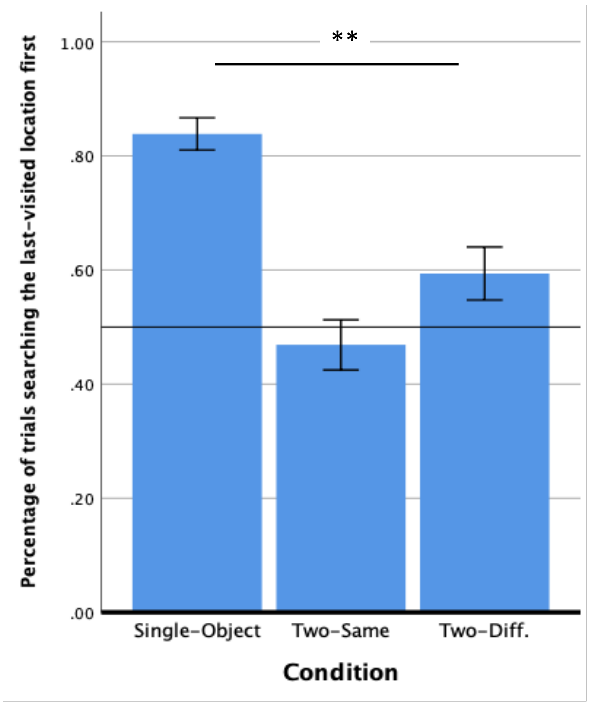


Figure 4-9 The percentage of trials that the monkeys first search the last-visited-location performance in familiarization phase.

Test Phase

For the same reason we explained in 4.2.2, given the definition of Second search, we considered the practice effect on a trial-by-trial basis in and only in *First search* and *Thorough search* analyses, but not in *Second search* analysis.

First search

A mixed ANOVA was used to analyse the Sameness (whether the two objects were of the same kind), the Temporal gap (whether there was a delay), and the Order effect (whether the monkey did the Delay block first or the Non-Delay block first), whilst taking the practice effect into account as a co-variate. In general, the monkeys were significantly more likely to search the last visited location first when the food items were different (the main effect of Sameness: $F(1,11) = 48.969$, $p < 0.001$, $\eta_p^2 = 0.860$. And they were marginally significantly more likely to do so when there was a temporal gap (the main effect of Temporal Gap: $F(1, 8) = 5.134$, $p = 0.053$, $\eta_p^2 = 0.391$). In fact, the significant interaction between Sameness and Temporal Gap ($F(1,11) = 6.710$, $p = 0.032$, $\eta_p^2 = 0.456$) suggested a moderation effect: monkeys showed a significant preference towards the last visited location in the Same trials than in the Different trials only when there was not a delay (Non-Delay: LSD(Same-Different) = 5.208, $p < 0.001$; Delay: LSD(Same-Different) = 1.083, $p = 0.322$). Besides, no effect involving Order was significant (Order X Sameness X Temporal gap: $F(1,11) = 3.260$, $p = 0.109$, $\eta_p^2 = 0.289$; Order X Sameness: $F(1,11) = 1.136$, $p = 0.318$, $\eta_p^2 = 0.124$; Order X Temporal gap: $F(1,11) = 1.975$, $p = 0.198$, $\eta_p^2 = 0.198$; Order: $F(1,11) = 0.520$, $p = 0.491$, $\eta_p^2 = 0.061$). None of the effects or interactions involving practice effect was significant (Practice: $F(19,209) = 1.486$, $p = 0.093$, $\eta_p^2 = 0.119$; Practice * Sameness, $F(19,209) = 1.164$, $p = 0.291$, $\eta_p^2 = 0.096$; Practice * Delay: $F(19,209) = 0.721$, $p = 0.795$, $\eta_p^2 = 0.062$; Practice * Sameness * Delay: $F(19,209) = 0.517$, $p = 0.953$, $\eta_p^2 = 0.045$). There is no evidence suggesting that the monkeys preference towards the last-visited location changed over trials.

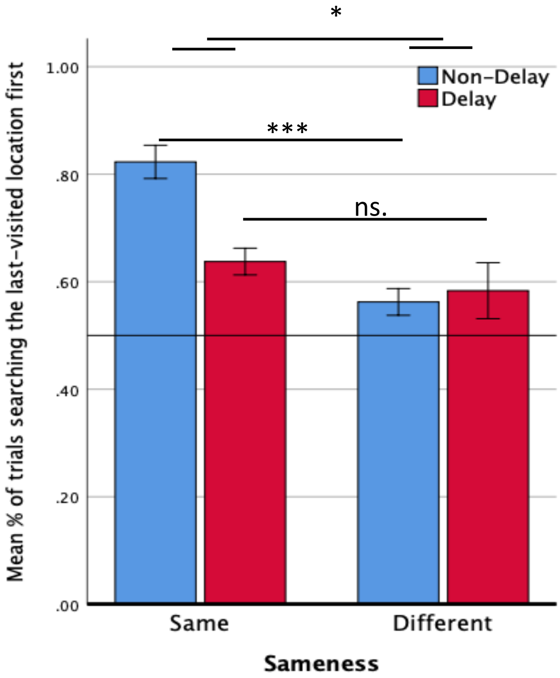


Figure 4-10 The squirrel monkeys' performance of "first search". The "first search" measures which location a monkey searched in a trial.

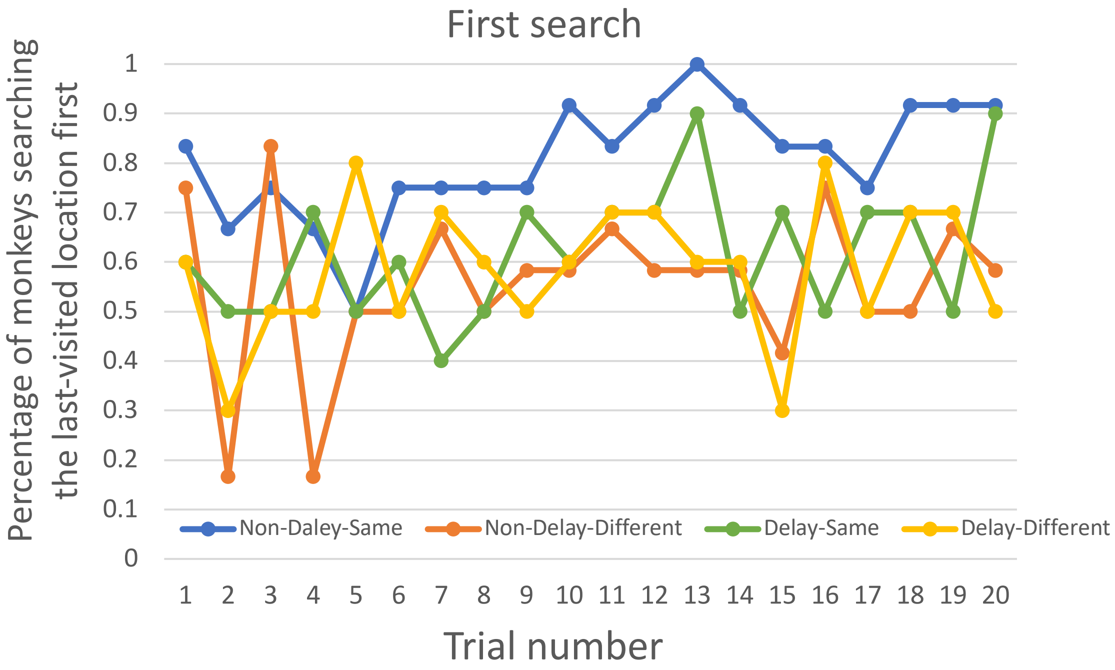


Figure 4-11 The squirrel monkeys' "first search" performance across trials.

Second Search

Again, we coded the "second search" (searching another location after retrieval from one location). A similar 2-by-2-by-2 ANOVA revealed no significant interaction between Sameness and the Temporal gap (F(1, 11) = 2.826, p = 0.131, $\eta_p^2 = 0.261$), no significant main effect on Sameness (F(1, 8) = 0.131, p = 0.727, $\eta_p^2 = 0.016$), and no significant main effect on Temporal Gap (F(1, 8) = 2.456, p = 0.156, $\eta_p^2 = 0.235$). All effects involving Order were not significant except for the main effect of Order (Order: F(1,11) = 7.236, p = 0.027, $\eta_p^2 = 0.475$; Order X Sameness X Temporal Gap: F(1,11) = 0.097, p = 0.736, $\eta_p^2 = 0.012$; Order X Temporal Gap: F(1,11) = 3.931, p = 0.083, $\eta_p^2 = 0.329$; Order X Sameness: F(1,11) = 1.546, p = 0.249, $\eta_p^2 = 0.016$). The main effect of order may suggest that in general the monkeys who experience the delay trials first would be more likely to search after retrieval. One potential explanation is that the monkeys focused better when starting with the longer version of trials (i.e., the delay trials), while they were more likely to lose interest if they got used to a shorter version (i.e., the non-delay trials) first before they had to wait for the additional three seconds during the longer version. This explanation could be supported by the similar findings in the Thorough search analysis below (i.e., only the main effect of Order was significant). We refrain from making further interpretations about this main effect, for three reasons: (1) the Order is a between subject factor, (2) both groups of monkeys had taken both the delay and non-delay trials, and (3) the Order did not (significantly) interact with any of the other factors.

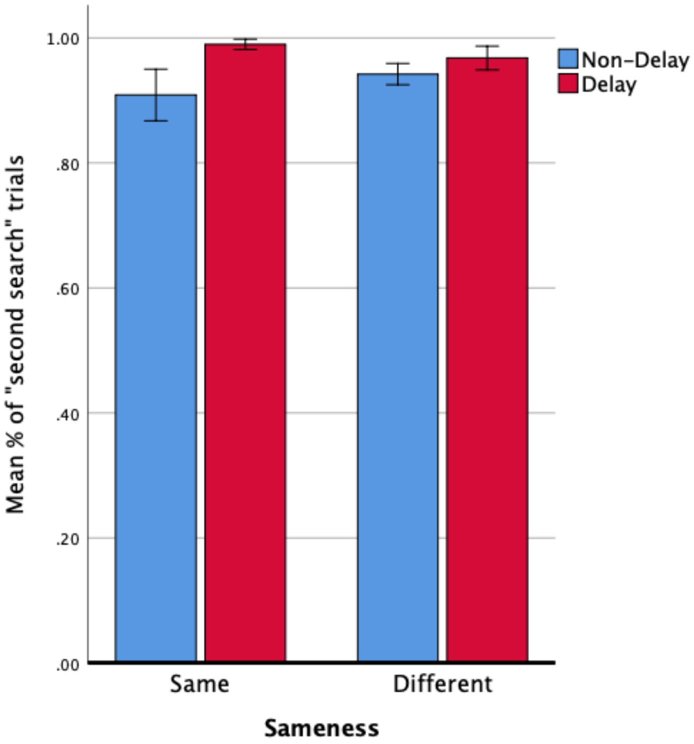


Figure 4-12 The squirrel monkeys' "second search" performance across conditions. The "second search" measures whether or not the monkey continued to search the other location after successfully retrieving the food from the last-visited-location.

Thorough search

We also coded the "thorough search" analysing the percentage of trials a monkey searched both locations regardless of the order. The ANOVA revealed no significant interaction between Sameness and the Temporal gap (F(1, 11) = 3.181, p = 0.112, η_p^2 = 0.285), no significant main effect of Sameness (F(1, 11) = 0.024, p = 0.881, η_p^2 = 0.003), and no significant main effect of Temporal Gap (F(1, 11) = 2.458, p = 0.156, η_p^2 = 0.235). All effects involving Order were not significant except for the main effect of Order (Order: F(1,11) = 8.571, p = 0.019, η_p^2 = 0.517; Order X Sameness X Temporal Gap: F(1,11) = 0.471, p = 0.512, η_p^2 = 0.056; Order X Temporal Gap: F(1,11) = 3.602, p = 0.094, η_p^2 = 0.310; Order X Sameness: F(1,11) = 0.597, p = 0.462, η_p^2 = 0.069). For similar reasons as above, we refrained from making further interpretations based on this main effect of Order. Besides, none of the effects or interactions involving practice effect was significant (Practice: F(19,209) = 0.864, p = 0.628, η_p^2 = 0.073; Practice * Sameness, F(19,209) = 1.126, p = 0.327, η_p^2 = 0.093; Practice * Delay: F(19,209) = 0.761, p = 0.752, η_p^2 = 0.065; Practice *

Sameness * Delay: $F(19,209) = 0.562$, $p = 0.929$, $\eta_p^2 = 0.049$). There is no evidence suggesting that the monkeys preference towards the last-visited location changed over trials.

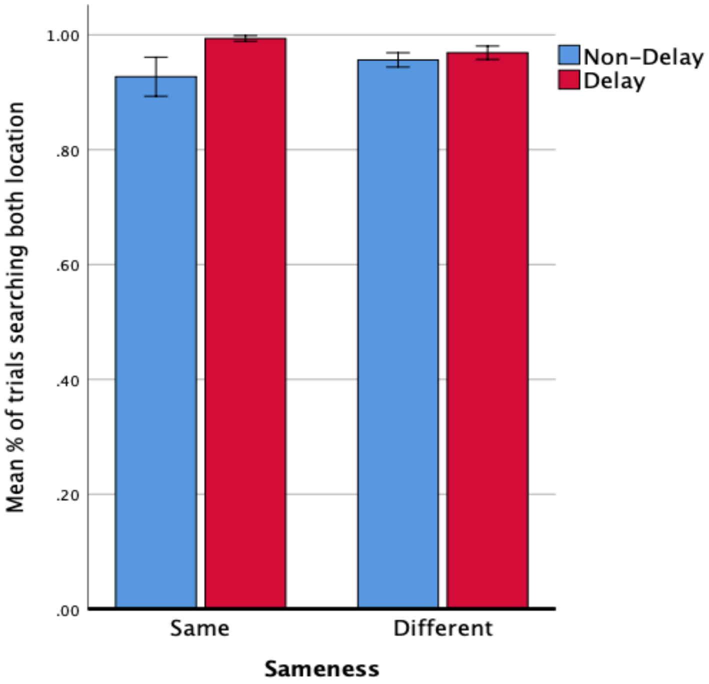


Figure 4-13 The squirrel monkeys' "thorough search" performance across conditions. The thorough search measures whether a monkey searched both two locations in a trial.

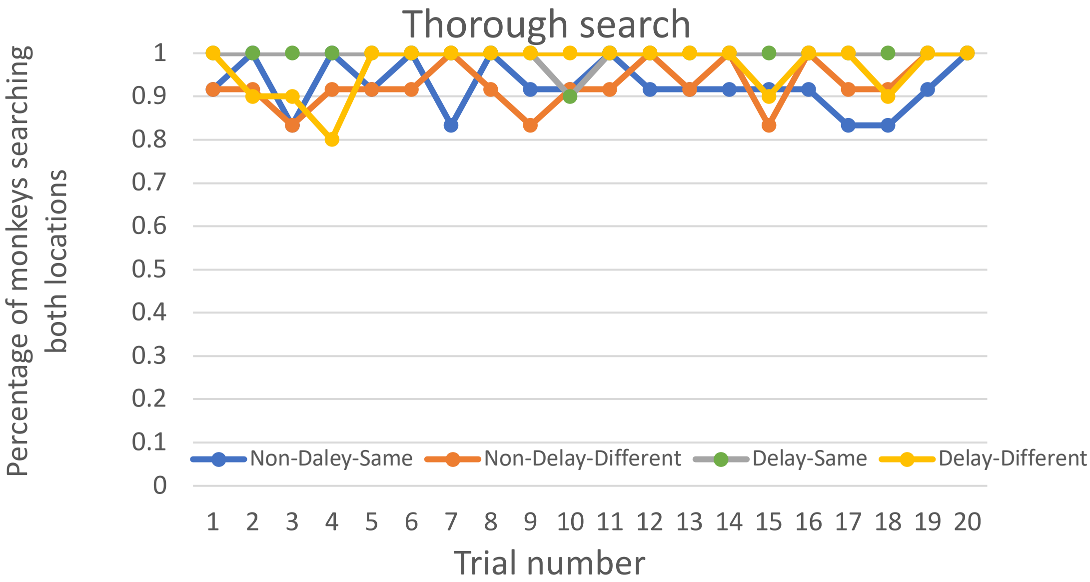


Figure 4-14 The squirrel monkeys' "first search" performance across trials.

4.3.3 Summary of results

Through a two-location paradigm, we investigated the squirrel monkeys' ability to individuate objects by property/kind. Our findings revealed the possibility that squirrel monkeys may possess property/kind individuation competence. When the two objects moved along a spatiotemporally continuous path (in Non-Delay conditions), the monkeys tended to regard the two objects of the same kind as a single object, and therefore searched preferentially the last-visited location. However, when it was two different objects moving on such a continuous trajectory path, they treated them as two, and searched equally between the two locations.

Hence, the squirrel monkeys in our study demonstrated their sensitivity to property/kind information when individuating objects. And they seemed to encode the property/kind information even when there was a conflicting spatiotemporal cue, without showing a "tunnel effect" or the spatiotemporal bias reported in rhesus macaques and humans (Bonatti et al., 2002; Burke, 1952; Flombaum et al., 2004, 2012; Scholl & Pylyshyn, 1999; Wilcox & Baillargeon, 1998a).

Interestingly, they showed no preference towards either location when there *was* a temporal gap between the movement of the two objects, regardless of whether they were the same or not. There are two possible explanations. On the one hand, it is possible that the monkeys lost track of the whole trial during the three-second gap, so that they just chose randomly between the two locations afterwards. If this were the case, which would be difficult to determine in our study, the monkeys' performance in the Delay trials would not relate to their object individuation ability. On the other hand, if the monkeys did follow through the whole event and made their choice accordingly, this result would suggest that the monkeys tended to believe there were two objects in both types of the Delay trial when there was a temporal gap between the occlusion of the first object and the emergence of the second object. Such phenomenon, although it looks very different from the "tunnel effect", can be explained by the Object File theory within the framework of the Carey-Xu account: when an individual perceives a temporal gap during the movement, a second object file would be opened on the basis of such discontinuity, regardless of whether the two objects were different or not (Carey & Xu, 2001). Taken together with the finding in Non-Delay-Different trials, this would mean that the squirrel monkeys encoded the property/kind information. These results are also compatible with the OF+PR model (Stavans et al., 2019) which emphasises the resolution of disagreement between systems encoding spatiotemporal information and property/kind information (for the potential disagreement, see Table 4-6). We will further discuss this in section 4.4 and chapter 6.

The findings regarding the temporal gap may also be able to shed some lights on our previous discussion (section 4.2.3) about the attentional effect interpretation. Is the monkeys' performance difference across conditions (in the spatiotemporal task) truly based on the expectation of object(s), or is it a result of an attentional bias that made the monkeys attend to both locations more equally because of the discontinuity? On the one hand, the movement path in the Non-Delay conditions of this experiment was always continuous. Thus, (dis)continuity would not make the monkeys attend differently between the locations. The more plausible explanation of our results is expectation-grounded. On the other hand, it is still difficult to completely rule out the possibility of an attentional effect in the Delay conditions, when there were temporal gaps involved in the event. For example, it is possible that the monkeys chose randomly in both the Non-Delay-Same and Non-Delay-Different conditions simply because their attention was divided between the two locations noticing the temporal gap, or as we discussed earlier, the monkeys completely lost track of the whole event. Again, we propose that it would be useful for future studies to address this issue with new techniques or paradigms.

In addition, our findings in the current experiment again revealed "ceiling effects" on the "second search" and the "total search" measures. A critical difference between this and the earlier spatiotemporal experiment in chapter 2 shall be noticed: in the continuous trial of the spatiotemporal task, there was only one piece of food hidden in one of the locations, whereas in the property/kind task, the two locations were always baited regardless of the conditions. Namely, in the current task, the monkeys were rewarded for searching twice. This might explain why we found no significant difference on these two measures comparable to those we found on the "first search" performance. However, we didn't find evidence to support this explanation, given the practice effect was not significant.

4.4 Discussion

In this chapter, we tested the squirrel monkeys with a two-barrier paradigm instead of the one-container "Magic box" task in chapter 2, and accordingly included a preference measure ("first search") rather than just the total search measure (e.g., "searching bouts/duration" in chapter 2 or "second/thorough search" in the current chapter). Given these squirrel monkeys' failure in chapter 2, we expected this adjusted combination of apparatus, paradigm, and measures could be more sensitive to test squirrel monkeys' object individuation abilities.

Our results in this chapter based on the "first search" measure suggest squirrel monkeys may also be able to use spatiotemporal and property/kind information to individuate objects, as capuchin monkeys, macaques, apes and twelve-month old human infants can do (Kersken et al., 2020; Mendes et al., 2008; Santos et al., 2002; Xu & Carey, 1996). These findings provide another piece of evidence that non- or pre- linguistic individuals are able to individuate objects by property/kind, which again challenges the strong version of Language Gain Hypothesis, the advocates of which argue that successful individuation by property/kind relies on the acquisition of language (Xu, 2002, 2007; Xu & Carey, 1996). Our results adding to the existing ones suggest that language may not be necessary for property/kind object individuation. At least, some nonhuman animals as well as preverbal human infants have been reported to pass the property/kind tasks. Though it is still possible that language plays a unique role in human development regarding object individuation, existing literatures in this field imply that the evolutionary origin of the ability to individuate objects using property and kind information is earlier than the strong version of the Language Gain hypothesis suggests.

Moreover, the current findings also fail to support the Carey-Xu account that Xu and Carey proposed to explain the developmental gap found in human infants between 10- and 12-month (Carey & Xu, 2001; Xu & Carey, 1996). According to this account, it is natural to expect a similar gap in at least some species. However, in all the comparative literature we reviewed in chapter 1, no such gap has been found. All the nonhuman animal species tested so far can use both spatiotemporal and property/kind information to individuate objects (Bräuer & Call, 2011; Fontanari et al., 2011, 2014; Kersken et al., 2020; Mendes et al., 2008; Santos et al., 2002). Our current findings with squirrel monkeys are in line with the existing literature. In the two-barrier tasks, squirrel monkeys showed sensitivity to both spatiotemporal and property/kind information of objects.

Furthermore, a more complex, broader picture appears, if we take both the “second search” and “first search” measures into account in comparison with our previous findings in chapter 2 with the “magic box” tasks. First, the squirrel monkeys failed the magic box experiment in chapter 2, while both capuchin monkeys, macaques, apes, and human infants have shown competence on similar tasks. Considering the “second search” and “thorough search” results in this chapter, it seems that the squirrel monkeys did have a tendency to search in both locations. Searching to both locations was also high in the single object familiarization sessions. And this redundant searching appeared to be a reliable pattern in not only the two-location paradigm we used in this chapter, but also the single-location paradigm in chapter 2. In the “Magic box” experiment, the squirrel monkeys also searched averagely more than once, even in the one-object conditions.

Why would squirrel monkeys act in this way? On the one hand, one of our aims was to make the apparatus appropriate and interesting to the monkeys (for example, the action of knocking down the semi-cylinders). Some monkeys may have found it so attractive that they wanted to interact with the apparatus even after they thoroughly searched both locations and retrieved all the food items. On the other hand, squirrel monkeys are not known as a species with strong cognitive control abilities (Fragaszy, 1985; MacLean et al., 2014). Thus, it may have been natural for them to launch an action with such low-cost and potential gain (the chance of finding more food rewards). As a consequence, measures based on frequency or duration of responses might not be the best option to test squirrel monkeys. Instead, we suggest that future research on this species should consider preference measures, such as the “first search” behaviour used in the present study.

Taken these two chapters together, our results present a mixture of positive and negative evidence for spatiotemporal and property/kind individuation in squirrel monkeys. On the one hand, the first search measurement suggest that they are sensitive to the probability of there being two objects vs one when shown property/kind evidence. On the other hand, they show a tendency to second search and search thoroughly beyond the individuation information available to them. Maybe it could be explained by a poorer cognitive control, as some literature suggests (Addessi et al., 2013; Anderson et al., 2010; Evans et al., 2012). Yet this assumption is not supported by our findings in chapter 3, in which the squirrel monkeys, compared to capuchin monkeys, didn't react faster or more redundantly when the stimulus (i.e., the cricket) was directly accessible to them. Or maybe the squirrel monkeys' object representation is relatively more fragile compared to other primate species tested. It may be necessary to move beyond the dichotomic question of whether a species demonstrates or not the different types of individuation into a more nuanced approach where we try to understand the different ways and extent to which they make use of the different types of information.

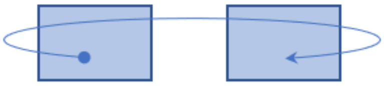
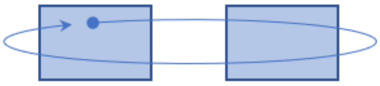
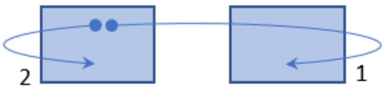
Another issue we would like to highlight here is the temporal aspect of spatiotemporal individuation. In the past two or three decades of individuation studies, most research has been focusing on property/kind individuation. In contrast, spatiotemporal individuation is usually taken as a basic ability that emerges early in both ontogeny and phylogeny. In addition, in these studies, spatial information is usually the predominant aspect for experimenters when designing different conditions, whereas temporal information is neglected (i.e. not manipulated or even described in the experiment). For instance, in the seminal articles assessing human infants' understanding of continuity principles, the most useful feature across conditions is the connecting trajectory between locations, i.e., spatial discrepancies (Spelke et al., 1995; Xu & Carey, 1996)[3]. The "tunnel effect" studies did consider the temporal delay during the movement of objects in occlusion events (Burke, 1952; Flombaum et al., 2004, 2012; Flombaum & Scholl, 2004; Scholl & Pylyshyn, 1999). Yet, the temporal information was always confounded with property/kind information. For instance, when an object A moves into the tunnel and an object B moves out and there is a delay, it is natural for a subject to think of two objects since both temporal and property/kind information suggests so. In contrast, if there is no delay, the spatiotemporal information suggests only one object whereas the property/kind information suggests two. It is notable that in neither condition were two similar objects presented following a temporal gap. Our current study tried to incorporate the temporal aspect into the object individuation tasks. In the spatiotemporal tasks, we designed a condition (Discontinuous B), in which an individual succeeds only if the temporal information is encoded as well. The results suggest that the

---

[3] Spelke and collaborators (1995) actually tested the 4-month-olds with a task in which object motion behind a barrier was smooth (no-delay) or not (delay). Yet their findings suggest younger infants didn't understand the so-called "smoothness" principle. Namely, younger infants encoded the spatial gap but not the temporal gap.

squirrel monkeys can take both spatial and temporal information into account (see Table 4-5 for a breakdown view).

Table 4-5 Breakdown by conditions of the two-barrier spatiotemporal task

Condition	Movement path	Object number	Spatial continuity	Temporal continuity	Monkeys' expectation
Cont. A		1	Yes	Yes	1
Cont. B		1	Yes	Yes	1
DisC. A		2	No	No	2
DisC. B		2	Yes	No	2

Note: Cont. = continuous; DisC. = discontinuous. The number in the diagram indicates the order of moving events (e.g., for DisC. A, the first food item comes out from the left barrier, goes back, and then the second food item appears and disappears on the right side.)

The food item starts from either the left (as illustrated above) or the right with counterbalanced order.

In the property/kind tasks, we adapted the design of the tunnel effect studies and manipulated the temporal delay of object movement behind the barriers. The results were unexpected (see Table 4-6 for a breakdown view of what each source of information indicates and how monkeys reacted). On the one hand, the squirrel monkeys first searched equally between the two locations when the temporal information indicated two potential objects, but the property/kind information indicated a single object (Delay-Same trials). On the other hand, the squirrel monkeys first searched equally between the two locations when the spatiotemporal information indicated one object, but the property/kind information indicated two objects (Non-Delay-Different trials). The stronger interpretation is that in both cases the monkeys assumed two objects. Namely, they encoded the temporal information in the former case, and encoded the property/kind information without showing the tunnel effect in the latter case. The interpretation for the former case is in line with the so-called spatiotemporal primacy. The primacy of spatiotemporal information (or spatiotemporal bias) is a major signature limit that supports the Carey-Xu account. According to the account, the first system indexes bounded physical objects (“Spelke objects”) and store/update the spatiotemporal information in symbolic representations (object-files). Therefore, featural or kind information will be ignored by this system. However, the latter finding of our experiment suggest that the property/kind information was processed and even overrode the spatiotemporal information during object individuation. From this point of view, the OF+PR model may provide a better framework to interpret these findings. This model emphasizes how the disagreement between systems is resolved rather than whether or not an individual is equipped with both systems. From this perspective, our results suggest that the squirrel monkeys resolved both types of disagreement, either when temporal information or property/kind information indicated more objects (see Table 4-6).

However, there is also another possible interpretation. When the spatiotemporal and property/kind information contradict with each other, the spatiotemporal information does not always prevail as the Carey-Xu account predicts. The disagreement may remain unresolved, leading to a random search between the two candidate locations. This interpretation is only partly compatible with the OF+PR model, given “whether there is a food item behind the first-visited location or not” is a “qualitative” question (for visualization, see Figure 6-1 C) and such kind disagreement should be resolvable according to the OF+PR model. Thus, neither theoretical framework is perfectly in line with this interpretation of our findings. Although it cannot be completely ruled out, future studies will be needed to test this and the former interpretations.

Table 4-6 The prediction based on temporal vs. property/kind information and squirrel monkeys’ performance on two-barrier property/kind individuation task across conditions

Condition		Temporal information	
		Non-Delay (1-object)	Delay (2-object)
Property/kind information	Same (1-object)	1-object	2-object
	Different (2-object)	2-object	2-object

Note: in the Delay-Same condition the information is actually ambiguous: it can be either two objects of the same kind, or the same single object that paused.

Taken together, our experiments reveal object individuation abilities in another new world monkey species, the squirrel monkey, and contribute to a fuller evolutionary picture of such abilities. Our findings, along with those with other nonhuman primate species, cast further doubt on the strong version of Xu and Carey’s Language Gain Hypothesis that property/kind object individuation requires language. Furthermore, squirrel monkeys were apparently encoding property/kind information even when it was conflicting with the spatiotemporal information. This finding would be compatible with the recent OF+PR model, which emphasises the process of resolving the disagreement between systems (Stavans et al., 2019), challenging the Xu-Carey account, which underscores the primacy of spatiotemporal information (Carey & Xu, 2001).

In the next chapter we will move from objects to agents in our exploration of core representational systems of knowledge.

Chapter 5 Capuchin monkeys represent hidden causal agent: Hidden agent experiment

5.1 Introduction

In the previous chapter, we focused on the monkeys' object representations, revealing that with the proper paradigm, measure, and experimental set-up, squirrel monkeys have shown some sensitivity to both spatio-temporal and property-kind information when individuating objects, though it is arguable that the squirrel monkeys may not perform the task or represent the object on a comparable level as apes, rhesus macaques, and capuchin monkeys do. These findings support the idea that monkeys share a core knowledge system for objects with human beings. The current chapter aims to extend the study by exploring another core knowledge system, the one representing agents.

Reasoning about the existence of a causal agent from observing the effects of its actions is a useful capacity: we can use this skill to avoid a lurking predator, or to track an unseen companion in a collaborative activity. Evidence from developmental psychology suggests that this skill emerges early in human development: babies can infer the existence of an agent based on the movement of an inert object in the first year of life (Saxe et al., 2005, 2007; Wu et al., 2016). However, in comparative psychology whether and to what extent non-human animals have similar abilities is a contentious issue.

Existing evidence suggests apes, as well as human infants, can infer the existence of a hidden physical object from observing effects by the object based on their understanding of causal principles and properties (Baillargeon, 1995; Call, 2004, 2007; Civelek et al., 2020; Hanus & Call, 2008; Hill, Collier-Baker, & Suddendorf, 2012; Phillips et al., 2009). However, when an agent acts as a cause, the predictive framework is more complex, because it involves understanding that agents carry out actions with intentions and goals (Baker, Saxe, & Tenenbaum, 2009; Biro & Leslie, 2007; Gergely & Csibra, 2003). For example, whereas a ripe apple will invariably tend to fall to the ground because of gravity along a straight trajectory, an agent can either pick it up, hold it, drop it, throw it, eat it, etc., with each option entailing a totally different pattern of movement. Understanding the causal power of agents is important, because agents are likely to be the ultimate causal source of many events: as self-propelled entities, they are capable of exerting force upon other objects and making them move or change state.

The theory of core knowledge systems proposed that representing and inferring about agents involves a cognitive module that develops early in human infants (Spelke et al., 1992; Spelke & Kinzler, 2007). Human infants can differentiate agents from physical objects (Bonatti et al., 2002; Kuhlmeier et al., 2004; Surian & Caldi, 2010). They can make different movement predictions based on whether the target is an agent or an object, or infer whether the target is an agent or an object based on the movement pattern (e.g., Luo et al., 2009; Träuble et al., 2014). And they can also attribute a goal to an agent based on the action (Csibra, 2008; Gergely & Csibra, 2003; Liu, Ullman, Tenenbaum, & Spelke, 2017; Luo, 2011; Woodward, 1998). In these examples, human infants consider an agent as a self-propelled entity with an internal causal power and possibly the goal to make itself move or to start a physical event. It has also been reported that nonhuman animals may understand agents and their goal-directed actions (Call et al., 2004; Phillips et al., 2009; Wood, Glynn, & Hauser, 2008; Wood & Hauser, 2008). Thus, it is likely that a core system of agent representation is underlying these examples, when the agent itself is directly observable.

Moreover, developmental research indicates human infants can attribute an event to an agentive cause, either observed (e.g., when the infants watched a box collapse because of the touching of a hand) (Muentener & Carey, 2010) or unobserved (Saxe et al., 2005, 2007; Wu et al., 2016). For instance, when 7- and 10-month-old infants saw a beanbag appear as if thrown from behind a barrier, they expected to see a hand on that side, and were surprised if the hand was visible on the opposite side, suggesting the infants anticipated an agent as the unobserved initial causal source of the event (Saxe et al., 2005, 2007). No such anticipation was found when it was a self-propelled puppet instead of an inert beanbag that moved across the stage, nor when instead of a hand, an inanimate toy appeared on either the same side or the opposite side where the beanbag came out (Saxe et al., 2005). These results suggest that the 10-month-old infants were capable of inferring an unobservable causal source for the movement of an inanimate object, and, more specifically, the cause should be an agent located in a particular place (where he could be responsible for the event). A recent empirical study suggested New Caledonian crows show a similar ability to infer hidden agents as an event's cause (Taylor et al., 2012). The crows needed to retrieve food from a container in front of a hide, from which a hidden human could poke a stick towards the container. In the Human Causal Agent Condition, the crows saw an experimenter enter the hide, and leave after the poking event, whereas in the Unknown Causal Agent Condition, they saw the stick poking without witnessing the experimenter entering or exiting. The crows inspected the hiding more often, and even abandoned the food retrieval during the Unknown Agent condition. Based on these observation, the authors suggested the crows did so because they inferred and attributed the poking event to an unseen agent. However, critics argued the crows could have learned to anticipate the agent presence in the Human Causal Agent condition, which always preceded the Unknown Causal Agent condition (Boogert, Arbilly, Muth, & Seed, 2013; Dymond, Haselgrove, & McGregor, 2013). Another argument is that the unmatched features between conditions weaken the claim that crows can reason about hidden causal agents. For example, the number of human agent(s) involved in each condition was different. In the Human Causal Agent condition, there were two humans present. Both experimenter A and B entered the room, where A stayed in the corner and B entered the hiding. After the poking, B left the hiding place and the room first, and then A also left. In contrast, in the Unknown Causal Agent condition, only experimenter A was involved (entered, stayed in corner, and left after the poking event). It might have been distracting for a bird to see the experimenter B leave the hiding after the poking event in the Human Agent condition but not in the Unknown agent condition, which may cause the reduced tendency to inspect (see details and the authors' replies in, Boogert, Arbilly, Muth, & Seed, 2013; Dymond, Haselgrove, & McGregor, 2013; Taylor, Miller, & Gray, 2013a, 2013b).

To further investigate the evolutionary roots of this ability, our current study tested capuchin monkeys[4] with a new hidden agent paradigm applied with a balanced design. This group-

---

[4] We tested capuchin monkeys (but not squirrel monkeys) given previous literature has reported successful physical and psychological reasoning in capuchins (see above and also section 1.3 and 3.1). Moreover, capuchins have shown a higher level of attention maintenance compared to squirrel monkeys according to the literature and also our previous studies (e.g., comparing squirrel monkeys' performance in chapter 2 and capuchin monkeys' in Kersken et al., 2020). Therefore, we assumed that starting with capuchin monkeys would be more helpful to establish the research paradigm. However, we acknowledge that it would be of interest to test squirrel monkeys in a further study with an adapted paradigm.

living species has shown sophisticated understanding of the physical (e.g., object manipulation and tool using) (Fragaszy et al., 2004; Luncz et al., 2016; Ottoni & Izar, 2008; Seed & Byrne, 2010) and social world (e.g., gaze following, perspective, goal-directed action, intention, etc.) (Barnes, Hill, Langer, Martinez, & Santos, 2008; Drayton & Santos, 2014; Hare, Addessi, Call, Tomasello, & Visalberghi, 2003; Phillips et al., 2009), making them an interesting subject species for the research question of whether nonhuman primates can trace back the cause of an event to a hidden agent.

5.2 Methods

5.2.1 Participants

Capuchin monkeys (*Sapajus apella*) from two social groups (east and west groups) at the Living Links Research Centre at Edinburgh Zoo (for more details about the facility: Macdonald & Whiten, 2011) participated in the present experiment on a voluntary basis. All capuchins were free to enter the research cubicles to separately attend the experiment, and were also free to leave anytime when they wanted. The monkeys were not food or water deprived and only positive reinforcements were used. The study was granted ethical approval and supervised by the University Teaching and Research Ethics Committee of University of St Andrews.

5.2.2 Apparatus and set-ups

The apparatus consists of two food dispenser boxes, a trolley table with a sliding tray on it, two large screens fixed to the table, a portable screen, a cylinder, a ramp, and a block.

The basic set-up is shown in Figure 5-1. The monkey in the cubicle was facing the sliding tray on the table. On the tray, there were two polycarbonate boxes (19 × 9 × 15 cm). Their front openings (3 × 4 cm) were aligned to the access holes of the cubicle so that the monkeys could reach in to get food. The boxes also had rear openings (8 × 9 cm) for baiting or fake baiting, covered by two pieces of card paper to block the light. Specifically, except for the first familiarisation phase, each of the rear openings led to a secret small compartment, so that when the experimenter pretended to bait, the food item actually fell in the compartment, so that the monkeys could only find the pre-baited one. These rear openings were aligned to the holes on the hiding screens, which were covered by flap doors made of cardboard, so that the experimenter could bait (or pretend to bait) the boxes from behind the screens. The fixed hiding screens were attached to the table base so that they didn't move with the tray. Through the gap between the two hiding screens, the monkeys could see a cylinder and a ramp/block in the middle of the table. In the arbitrary control trials the cylinder was placed on the ramp/block so that it would roll/fall down after a shake of the table. In the agentive trials, the ramp/block was also present in the scene but only behind the cylinder, serving as a background to match the control trials. Behind all this there could be a portable screen with hooks, which was set up to occlude the experimenter or be removed to reveal him. In addition, a linen sheet hanging on the rear edge of the table always prevented the experimenter's body from being seen by the monkeys.

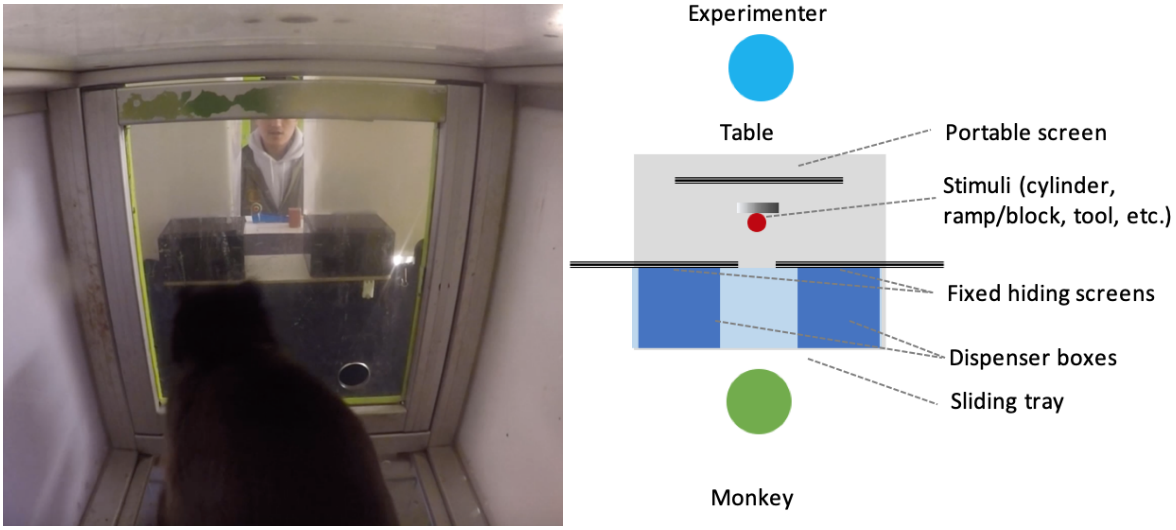


Figure 5-1 Left panel: the picture of the set-up from the monkey's perspective (without portable screen); right panel: plan view illustration of the set-ups.

The task involved monkeys searching for food in one of two boxes. First, they learned that the agent will always hide food in the box that he approached and stood behind. Then, the area behind the 2 boxes was occluded, and the monkeys had to infer behind which screen he was standing from a physical event (e.g., a cylinder being rolled across the table from behind the left screen). In the ***Agentive*** condition, the agent was the cause of the movement event. In the ***Arbitrary control*** condition, the movement of the cylinder was caused by the table shaking, after which the cylinder rolled down a ramp or off a platform (an L-shape block). If monkeys understand that agents caused the movement, they shall search behind the screen from which the movement originates, but only when there is no other causal explanation for the direction of the movement (i.e., in the causal agent condition but not the arbitrary control condition).

5.2.3 Procedure

Step 1: Familiarization tasks

We conducted 3 familiarisation tasks to make sure that (1) the monkeys had no difficulties in interacting with the apparatus, (2) they could understand the experimenter's baiting action and could use it as cue to find the hidden food, and (3) they understood that the experimenter was baiting when he moved behind one of the screens although the baiting action was not directly visible.

Each familiarisation task followed the procedure explained below. The experimenter introduced the monkey into the cubicle, went behind the table, drew the monkey's attention by opening the food container and retrieved one piece of food that held in his hand in the middle point behind the table. The experimenter then moved to one side of the table towards one of the boxes with the food. After the food item was inserted into one of the boxes, the experimenter pushed the tray forward so that the monkey could make a choice by reaching into (or touching) one of the boxes. The experimenter reappeared in the middle after the choice. Then he moved to the monkey's side of the table, blocking the monkey's sight whilst reaching into both boxes at the same time to check for any leftover food. This would be handed to the monkey if it had succeeded in this trial but failed to retrieve the food on its own, or would be recycled for later use if the monkey chose the wrong box. Meanwhile, the food item for the next trial was surreptitiously put inside one of the boxes.

In **familiarisation task I (Fa1)**, the lids of the boxes and the screens were both removed so that the baiting and the final location of the food item were fully visible to the monkeys.

In **familiarisation task II (Fa2)**, there were lids on the boxes but still no screens, so that the monkeys would saw how the experimenter inserted the food into one of the boxes (the food item actually fell into the secret compartment). The monkeys then were allowed to search by themselves accordingly.

In **familiarisation task III (Fa3)**, there were both lids and fixed screens so that the monkeys could only saw that the experimenter hide behind one of the screens. In familiarisation IIIa trials (F3a), the experimenter pretended to bait in a way similar to that in Fa2, but through the flap doors on the screens (which were small so that these doors were occluded by the boxes and the monkeys could not see the experimenter and the baiting actions through them), while in the familiarisation IIIb (Fa3b), the experimenter just waited for approximately 2 seconds as if he was baiting invisibly before pushing the tray forwards, without doing any baiting action. Thus, the monkeys needed to infer the location of the food according to their previous experience, namely tracing the position of the experimenter to find the hidden food, given that the food was baited by the experimenter and therefore now it should be found on the side where the experimenter went to.

Each task contained 10 trials per session (5 trials for each side in a counterbalanced order). Monkeys that found the correct location in 8 out of 10 trials proceeded to the next stage. If an animal failed to meet this criterion in a session, the experimenter carried out the same task again until the monkey passed, or the number of sessions on this task for this subject added up to three. In the latter case we dropped this subject from the study.

Step 2: Test Tasks

There were 4 pairs of conditions, each pair included an agentive task and an arbitrary control task. In the agentive tasks the events were started by the hidden experimenter, while in the control tasks the similarly looking events were started by a shake of the table, therefore providing an ambiguous clue as to where the experimenter was located. It is noteworthy that in either condition, the event was performed by the experimenter from under the table. In an agentive trial, the experimenter crouched underneath the table and raised his hand into the occlusion of one of the screens to surreptitiously manipulate the tool or the stimulus. Whereas in an arbitrary control trial, the experiment under the table held the column of the projector trolley table and shook it briefly. Each task contained 2 trials with the factor of side/direction counterbalanced. The procedures of the tasks were the same except for the demonstration parts.

Before each trial, the experimenter pre-baited one of the boxes either before the monkey entered the cubicle (in the first trial of a session) or at the end of the last trial when the experimenter checked the boxes (see details below). In the agentive trials, the food item was on the side where the causal agent was supposed to hide, indicated by the cueing event (e.g., the side that the stick or the cylinder came from). In the arbitrary trials, the food item was baited in the box on the equivalent side, relative to the cylinder's moving trajectory which was similar to that in the agentive trials (for example, the food item was baited in the left when the cylinder rolled to the right in the pushing and the pushing-ramp trials, see the first row in Figure 5-2).

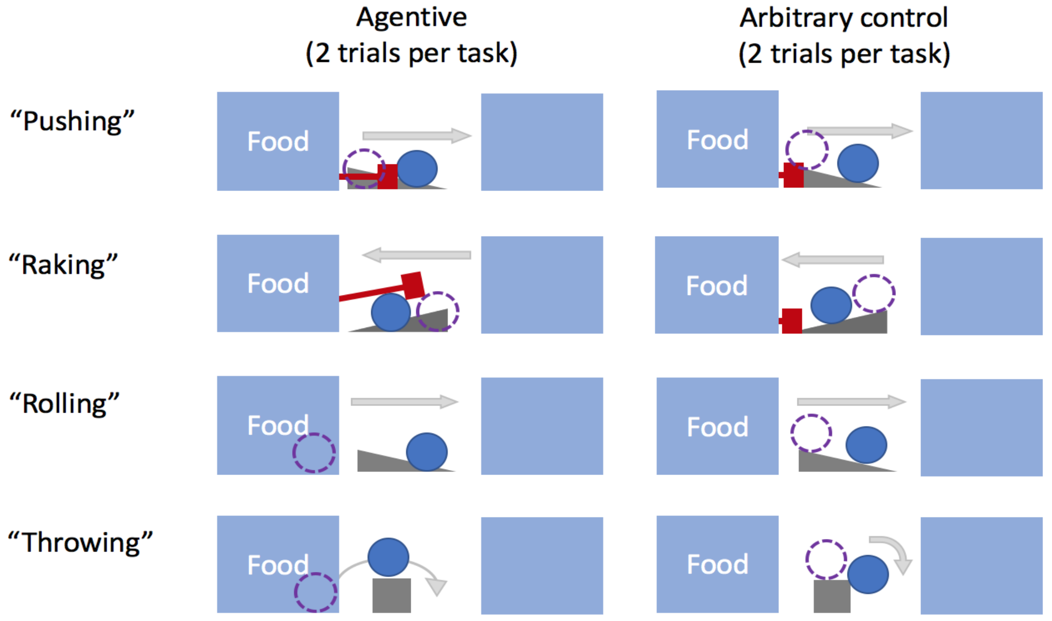
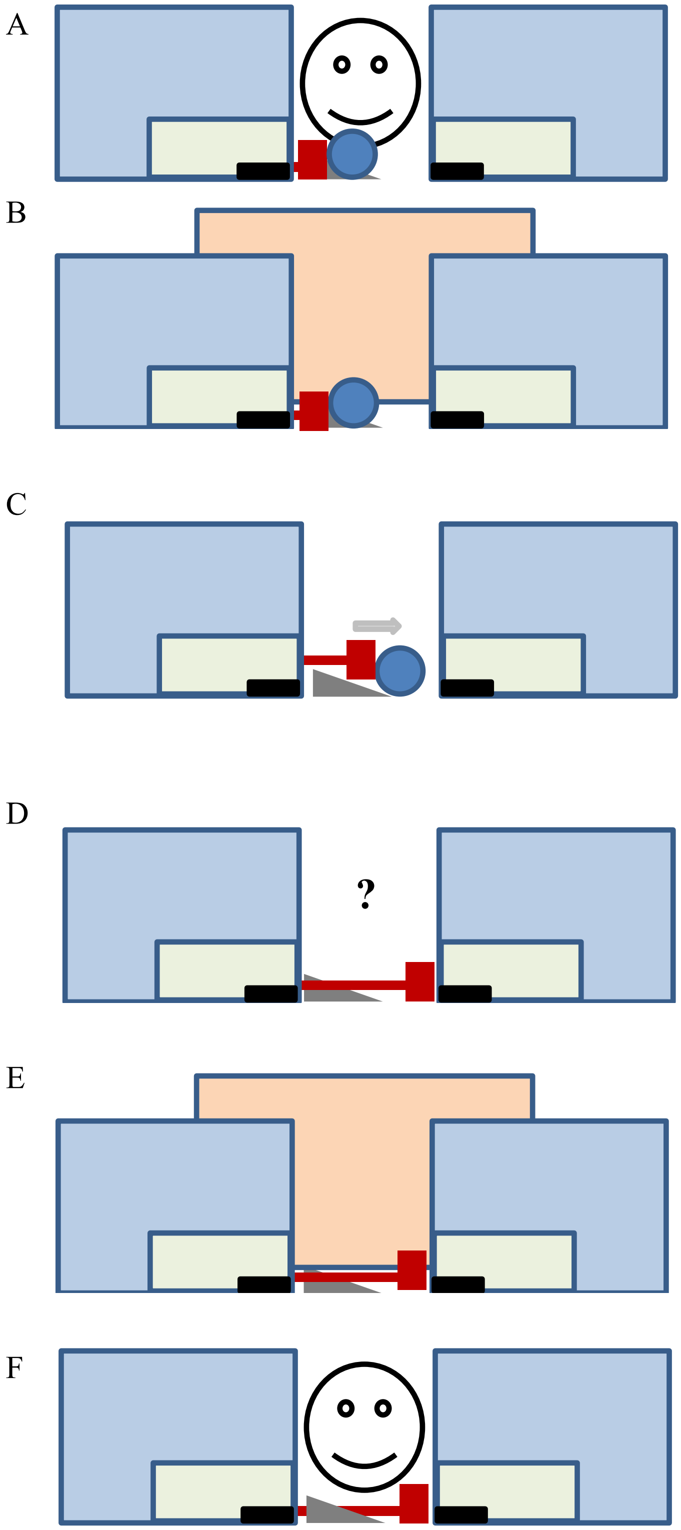


Figure 5-2 Illustration of the different conditions, only taking the trials that the food was hidden on the left as examples. The two big rectangles on both sides represent the two hiding screens, between which is the 'stage' and stimuli. The circle represents the cylinder and the dash circle marks the starting point of the cylinder's movement; the triangle in the middle represents the ramp; the small square in the middle represents the block; the hammer-like figure represents the tool; the arrows indicate the movement pattern.



The monkey comes in the cubicle and sees the experimenter in the middle of the whole set-up

The experimenter covers himself with a portable screen, crouching down to hide under the table, and making a ‘click’ noise by opening the food container,

The hiding exprimenter presents a cylinder movement, the reason of which is dependent on the conditions (here the cylinder is pushed to the other screen with a stick).

The monkey faces a choice when the tray and boxes are slid forward so that he/she can reach into one of the boxes.

The portable screen is set up again, allowing the experimenter to come back without revealing the location he was supposed to be hiding behind.

The experimenter reappears in the middle, and hands over the prebaited food item if the monkey make the correct choice but failed to retrieve it.

Figure 5-3 the general procedure of the test trials, taking the agentive-pushing condition as an example.

The monkey was introduced and separated in the cubicle, getting a sunflower seed as reward. The experimenter moved back to his middle position behind the table and drew the monkey’s attention by waving a hand and calling its name. He then set up the portable screen to allow the experimenter moving behind one of the screens, while in fact he crouched to hide right in the middle under the table to avoid any unwitting cue of his side location. After that, the portable screen was put down, showing the monkeys that the experimenter was no longer visible and therefore had hidden somewhere. The experimenter then made a ‘click’ sound in the middle by opening a food container box, giving the monkey an illusion that he was going to bait. The hidden experimenter then triggered a demonstration of the cylinder’s movement depending on the condition (performing by the experimenter under the table), Then the tray along with the dispenser boxes on it was slid forward to the monkey so that it could make a choice.

After the choice or approximately 10 seconds later, the tray was pulled backwards and the portable screen was installed. The experimenter then came back to the middle, removed the portable screen, and went to the monkey’s side of the table to do the check and pre-baiting. He reached into both boxes from the rear openings simultaneously, pretending to check if there was any leftover. The leftover food would be given to the monkey if it failed to get the food after making the correct choice, or would be recycled if the monkey made the wrong choice. Meanwhile, the experimenter surreptitiously hid the food reward for the next trial in his closed fist and baited one of the boxes.

The monkeys’ reactions as well as the demonstration of the event were recorded with a video camera situated on the rear ceiling of the cubicle, filming from the monkey’s perspective (see figure 5-1). The experimenter watched the real time video via the Gopro mobile application on his cell phone remotely to monitor the monkey and the event whilst he was hidden under the table.

To assess inter-rater reliability, a naïve second coder scored a random sample of 20% of the trials. Fleiss’ kappa was calculated, revealing a high level of agreement on which side the monkeys chose to inspect first for the food: Kappa= 0.875, $p < 0.001$.

5.3 Results

Sixteen capuchin monkeys participated in the familiarization phase, including four familiarization tasks (Fa I, Fa II, Fa II a, and Fa IIIb). A participant could have up to three ten-trial sessions to meet the criterion, which is eight successes out of ten trials (i.e., $p < 0.05$ according to binomial test), so that it could proceed. Ten monkeys passed all the familiarization tasks and all of them entered the test phase. The details of attendance and performance can be found in Table 5-1.

Table 5-1 Capuchin monkey subjects' attendance and performance on Familiarization tasks

No.	Subject	Age (yrs)	Fa I Session	Fa II Session			Fa IIIa Session			Fa IIIb Session		
			1	1	2	3	1	2	3	1	2	3
1	**Kato**	12	10	9			10			8		
2	**Hazel**	4	10	9			9			8		
3	**Inti**	8	9	quit	9		7	8		9		
4	**Torres**	6	10	quit	10		7	10		quit	10	
5	**Pixie**	4	10	10			quit	7	10	7	9	
6	**Reuben**	7	10	10			10			10		
7	**Rufo**	8	9	6			quit	8		9		
8	**Ximo**	7	9	6	10		8			quit	9	
9	**Chico**	8	10	10			9			10		
10	**Alba**	6	10	10			10			9		
11	Junon*	17	9	7	9		7	9		7	5	quit
12	Luna	6	10	5	quit	quit						
13	Carloz	11	10							quit		
14	Flojo	6	10	6	10		quit	quit	quit			
15	Willow	5	10									
16	Mekoe	9	9									

Note: The subjects who passed the familiarization phase were marked in BOLD font. "Quit" means the subject complete less than 5 trials before asking for leaving or being seriously distracted in a session., Junon entered the test phase as she insisted, though she didn't pass Fa IIIb. Her data was excluded from the analysis.

A two-by-four repeated ANOVA was conducted to analyse the data considering the effect of Agency (Agent vs. Arbitrary Control), the effect of Movement Type, and their interaction. We found only a significant effect of Agency ($F(1,9) = 12.522$, $p = 0.006$, $\eta_p^2 = 0.582$).

The effect of Movement type (F(3, 27) = 2.418, p = 0.088, $\eta_p^2 = 0.212$) and the interaction (F(3, 27) = 1.161, p = 0.343, $\eta_p^2 = 0.114$) were not significant. The capuchin monkeys were more likely to select the "correct" location in the Agentive trials than in the Arbitrary control trials. Given that the interaction was not significant and that the number of trials in each movement type was relatively small, we combined the four types of movement in each condition in the following analysis.

We also compared the monkeys' performance with chance level in each condition with one-sample t-tests, assuming chance was an equal probability of choosing each side. However, monkeys selected the correct side significantly below chance in the Arbitrary/Control conditions, demonstrating a bias to select the side where the cylinder moved (t(9)=-3.354, p=0.008, d=-1.061). This bias disappeared in the Agentive trials, although choice did not go above chance level (t(9)=1.500, p=0.168, d=0.474).

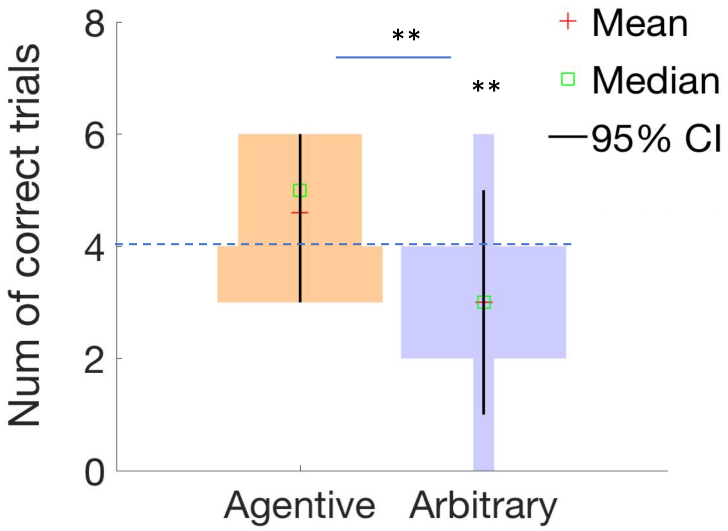


Figure 5-4 Number of trials in which the monkeys chose the correct side. The correct side was the side with food and the side that the experimenter SHOULD be in the agentive trials or the equivalent side in the arbitrary control trials. The orange represents the distribution of the correct trial numbers in agentive condition and the purple represents the distribution of number the correct trials in arbitrary control condition. The dash line represents chance level (50%). ** means p<0.01.

5.4 Discussion

We found the capuchin monkeys behaved differently when the path of an object's motion was consistent with it having been caused by an agent acting from a particular side, compared to control trials where the exact same movement occurred by an arbitrary cause not linked to any particular location of the agent. They used the movement of the object to infer the location of a hidden reward at higher levels in the agentive trials than in the control, non-agentive trials.

However, in the control non-agentive trials the monkeys showed a bias to select the box towards which the cylinder moved. Because three-quarters of the trials involved the object moving away from the baited location, this led to below chance performance in the control trials. This suggests the monkeys' choices were affected by two factors: the direction in which the cylinder moved, and the kind of movement event (agentive/control). On the one hand, the cylinder attracted the monkeys' attention towards the direction it moved. This explained why the monkeys responded at below-chance level in the arbitrary control trials, when the demonstration did not convey information about where the experimenter and the food should be. Our expectation was that the monkeys should be choosing at random, but this "direction of cylinder movement" bias, or the endpoint bias, led them to choose non-randomly. Nonetheless, the monkeys overcame this bias in the agentive trials when there was enough information for them to interpret the movement as provoked by an agent operating from one side, as shown by the significant difference we found between the performance on Agentive versus Arbitrary control trials.

With the different types of movement used in our experiment, it seems unlikely that the monkeys succeeded by using simple rules based on the direction of the movement and the location of the food (e.g., "follow/trace the movement to find the food") or the layout of the tool (e.g., "the experimenter and the food hid behind the side that the tool connected to"). In addition, the main perceptual features of the movement events in the control conditions were matched to each of the agentive conditions. The layout of the stimuli (the location of the tool, ramp, or block) and the direction of the cylinder's movement were matched as closely as possible.

Hence, our results seem to suggest that the capuchin monkeys made inferences about the location of the experimenter and where they should have placed the food when the movement pattern of the cylinder suggested an agentive force operating from one particular side. They did so by overcoming a bias towards the side where the cylinder moved when no information about the agent's location was available.

One may dispute the interpretation that the performance in Agentive trials was not above chance because of an endpoint bias. To further examine this interpretation, we reviewed the data of two particular pairs of conditions, "raking" and "rolling". In the Raking-Agentive trials, the stimulus was moving towards the location where the experimenter should be. The number of correct trials in this condition was almost higher than the chance level ($p = 0.057$, marginally significant), although we didn't find the significant effect of Agency in this pair of conditions (probably due to the small number of trials). Thus, it is presumable that with a larger number of trials and sample size, this comparison may reach significance. But on the other hand the number of correct trials in the Raking-Arbitrary condition was not higher than chance level, which would not support our assumption of the endpoint bias. Another observation and post-hoc analysis may help to solve this question: in the Rolling conditions, we found both a significant effect of Agency ($p = 0.003$) and the number of correct trials in Rolling-Arbitrary was significantly lower than chance level ($p = 0.006$). A possible explanation is that in the Rolling conditions, the cylinder rolled all over across the stage into the hiding location, leading the subject's attention to the opposite direction of the experimenter's plausible hiding location. Namely, the data was in line with our assumption of a strong end-point bias which was (partly) overcome in the Rolling-Agentive trials. In contrast, the cylinder normally stayed on the stage in the other types of trials. Thus, the endpoint bias might be much weaker in the other pairs of conditions, which is in line with the data (no difference between agentive and arbitrary control conditions and no difference from chance level, ps > 0.05).

Given the limited number of trials and relatively small sample size, we refrain from drawing a strong conclusion based on these post-hoc analyses. However, we suggest that these findings, though not decisive, imply the existence of an endpoint bias in our study. We would also suggest here some potential directions for follow-up studies to verify our interpretation of the current results: (1) expanding the number of trials or/and the number of condition(s) in which the stimulus moves towards the agent (or where the agent SHOULD be), like the Raking conditions here; (2) expanding the number of trials or/and the number of condition(s) in which the stimulus moves across the stage and disappears behind one of the barriers, like the Rolling conditions here. In other words, future studies may manipulate the direction of the stimulus (towards or from the agent's side) or/and the destination of the stimulus in order to better dissociate the endpoint bias and the agent-related reasoning.

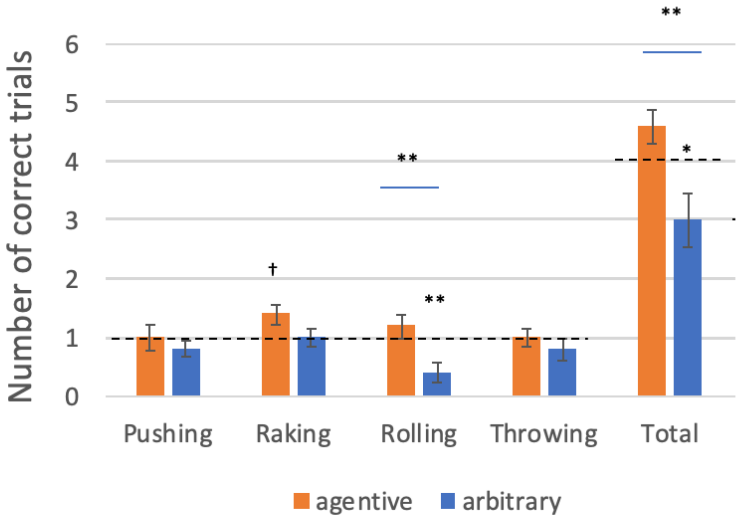


Figure 5-5 Number of trials in which the monkeys chose the correct side, broken down by movement types. The correct side was the side with food and the side that the experimenter SHOULD be in the agentive trials or the equivalent side in the arbitrary control trials. The dash line represents chance level (50%). ** means $p < 0.01$; * means $p < 0.05$; † means $p < 0.10$.

The competition between attending to the starting point and the endpoint of a movement is also evident in studies of human development. Research suggests human infants and children encode the starting point versus endpoint of an event differently depending on whether or not it involves an agentive cause (Lakusta & Carey, 2015; Lakusta & DiFabrizio, 2017). For example, in a series of dishabituation tasks (Lakusta & Carey, 2015), after being familiarized with an event involving the motion of an agentive object (e.g., a duck walks into a box), human infants spent longer time watching new events with changes in the endpoints rather than those with changes in the starting points, which indicates they were expecting any changes to occur in the starting point when an agent was involved. In contrast, this "endpoint bias" disappeared when the event involved non-agentive/inert object's movements (e.g., a leaf floats into the box), where infants attended more to changes in the starting points, implying they attended and reasoned about the cause of the movement when there was no agent in sight. Similarly in our results, capuchin monkeys seemed to prefer the endpoint when understanding the movement did not require an agent, while they shifted their interest towards the starting point when the source of the movement was most plausibly the hidden experimenter.

In future studies, it would be desirable to expand these findings, to explore if these expectations are unique to hidden agents. In the original experiments that inspired this study (Saxe et al., 2005, 2007), infants had expectations about the location of agents but not another inanimate objects, based on the path of the target object's movement. We would predict that, if the monkeys learned to associate the location of an object (instead of the agent in our study) with the location of food, they would not use the path of the target object' motion to infer the baited location in the same way that they did when their task was to locate an agent. But this hypothesis remains to be tested.

Taylor et al.'s crows' experiment that inspired the current study pointed out another approach. In their experiment, the critical manipulation was whether there was an agent leaving the hiding location where the poking started. The agent served as a "threat" and the bird would be at ease to approach the food only after seeing he leave. What would a bird do if it were an object (e.g., a trolley) instead of a human that leaves the hiding location (Boogert et al., 2013)? Will the bird keep vigilant, or will it behave as if the agent had left (Hidden Causal Agent condition)? In other words, will the bird attribute the poking to a leaving object as well as to a leaving agent? The answers to these questions may improve our knowledge of how crows understand and represent agents, especially in a causal context. This paradigm and its modifications could also be incorporated into our current task for capuchin monkeys, in a context in which the agent acted as a "helper" or "feeder". For example, the capuchin monkeys could be trained first to get food directly from an experimenter (most of the monkeys in the current research site already learned to do so), and then via a tunnel or access hole on the screen while the agent is not directly visible to them. In the test phase, after the agent disappears and an event is demonstrated on the stage, the monkey would be allowed to make a choice (similarly to the current study). Then the critical part comes: how frequently would the monkeys continue waiting (or begging) for, after either (1) the agent leaves, (2) an object (e.g., a chair) leaves (control condition), (3) a pause and nothing happens (baseline condition)? We expect future studies to consider these follow-up experiments. Some other discussions and limitations can be found in section 6.2.2.

In conclusion, our study provides preliminary evidence that capuchin monkeys have the ability to infer a hidden agent as the cause of a motion event, an ability that may be shared with human infants and (possibly) New Caledonian crows. These kinds of tasks with hidden agents simulate scenarios that may be highly relevant in the animals' natural environments and highlight the adaptive value of inferring the potential presence of an individual that is not yet visible (and who might be a friend, an enemy, or a predator) in a particular location from the causal structure of the events. Our findings with capuchin monkeys suggest that this highly adaptive ability has evolved in primates. However, a vast part of the broader evolutionary picture remains dim, especially in consideration of the evidence of the crow study (Taylor et al., 2012). Is it a convergent evolution in New Caledonian crows or an evolutionarily ancient adaptation that perhaps one can trace back to the early ancestor of vertebrates? Little is known and more species await to be tested to answer such questions. Nonetheless, our hidden agent paradigm provides a new tool (with some potential future modifications such as those mentioned above) for further research with other nonhuman primate species.

Chapter 6 General discussion

Representing objects and agents are basic cognitive abilities for human beings, emerging early in ontogeny (e.g., Saxe et al., 2005, 2007; Stavans et al., 2019; Xu & Carey, 1996). According to a core knowledge point of view, these representational systems are likely to have a long history in phylogeny, which means one can expect to find such abilities in nonhuman primates (e.g., Hauser & Spelke, 2004; Mendes et al., 2008; Santos et al., 2002; Taylor et al., 2012; Vallortigara, 2012). Moreover, if we agree with the presumption that these abilities rely on some core modules which are structurally similar across species, i.e., that human infants and nonhuman primates use similar systems to represent objects and agents, studying the performance of nonhuman primates may provide psychologists with a chance to review the theories and models that originally derived from developmental research from an evolutionary/comparative perspective (Hauser & Spelke, 2004; Weiss & Santos, 2006).

The current thesis tested some aspects of the object and agent representational abilities of two closely related new world monkey species, squirrel monkeys and capuchin monkeys. We focused on scenarios in which subjects needed to represent hidden entities, either objects or agents, based on various types of information. Specifically, for object representation, we examined whether squirrel monkeys were able to individuate objects based on property/kind information as well as spatiotemporal information, whereas for agent representation, we examined whether capuchin monkeys were able to infer the location of a hidden agent based on its actions and causal effects.

6.1 Squirrel monkeys individuate objects

6.1.1 Review and summary of results

The ability of object individuation enables an individual to track and identify objects based on the spatiotemporal and/or property/kind information. Initial research revealed that human infants can do the former before they can do the latter, which indicated a stage-like development that, it was suggested, might be related to language acquisition (Carey & Xu, 2001; Xu, 2002, 2007; Xu & Carey, 1996). Xu and Carey proposed a Carey-Xu model to explain these findings: (1) an early-developing object-file / object indexing system that processes spatiotemporal information and represents mere physical objects ("Spelke objects") and (2) a late-developing kind-based system that processes categorical information, and which highly relies on language (Carey & Xu, 2001; Xu, 1999; Xu & Carey, 1996; Xu et al., 2004). Note that according to this account, featural information alone is not sufficient to open a new object-file but can be bound to an existing object-file. Therefore, preverbal infants with only the first system are predicted to fail property/kind individuation tasks. However, using simplified tasks, research found evidence inconsistent with this Carey-Xu account, indicating that preverbal infants might be able to individuate objects by property/kind information at an earlier age than Xu and Carey suggested. Baillargeon and colleagues proposed a new model to reconcile these findings, in which an object-file system encodes both spatiotemporal and categorical information, whereas a physical reasoning system encodes featural information related to the event (Stavans et al., 2019). Critically, the OF and PR systems emerge, develop, and work in parallel rather than in succession, so that an individual may fail to individuate because of the disagreement between systems rather than the lack of one of the systems (Baillargeon et al., 2012; Carey & Xu, 2001; Stavans et al., 2019).

In this debate, one of the central issues is whether there is a stage-like "gap" between spatiotemporal and property/kind individuation. More specifically from a comparative perspective, Xu and Carey's two-system account predicts a species that passes the spatiotemporal individuation task but fails the property/kind one (Carey & Xu, 2001). The existing animal research clearly does not favour this account. All nonhuman primate species tested, including apes, rhesus macaques, capuchin monkeys, and even some other species including dogs and chicks, have been reported to individuate objects based on both spatiotemporal and property/kind information, thereby questioning the assumption that the latter type of individuation might be linked to language (e.g., Bräuer & Call, 2011; Fontanari et al., 2011, 2014; Kersken et al., 2020; Mendes et al., 2008; Santos et al., 2002). One may argue that this dissociation may only be found in ontogeny. For the nonhuman animal species tested above, there may be a timepoint in development in which an individual can individuate objects by spatiotemporal but not property/kind information. Current evidence in nonhuman primates cannot directly contradict nor support this assumption, since only adult individuals have been tested. However, the newly-hatched chick studies imply that these birds may process both individuation abilities at a very early stage of their lives, and it is unreasonable to posit a developmental gap here (Fontanari et al., 2011, 2014).

Furthermore, the Carey-Xu approach could also be defended (with a non-existent or more secondary role of language in the property/kind system) assuming that a possible explanation for these findings would be that the "gap" or the critical evolutionary timepoint separating both types of individuation is more ancient than researchers thought before. In other words, the ability that underlies property/kind individuation may be a primate primitive, which means all primate species should be able to pass the tasks. To examine this assumption (and maybe also the assumption above) and to further our understanding of this question, one approach is to expand the research to test more distant species. Thus, in the current thesis, we tested squirrel monkey.

The more recent model, proposed by Baillargeon and colleagues' OF+PR suggests that successful property/kind individuation depends on a suitable type of event and the task context (Stavans et al., 2019). Although this model is developmental only and does not address or speculate about the evolutionary origins of object individuation, one can try to apply it comparatively. For example, the magic box experiment with manual search measures is a widely used paradigm for testing nonhuman primates' object individuation abilities (e.g., Kersken et al., 2020; Mendes et al., 2008; Santos et al., 2002). The OF+PR model would predict that even younger infants (<10mo) would successfully individuate in this task, given that it involves more than one event (a containment event followed by a searching event) and the outcome from the two systems may differ qualitatively instead of quantitatively, which means the disagreement can be resolved. One could therefore assume that nonhuman primates would also be able to succeed in this type of task from the point of view of the OF+PR model.

The existing nonhuman primate studies using this paradigm appeared to support these presumptions (e.g., Kersken et al., 2020; Mendes et al., 2008; Santos et al., 2002). Especially considering the success in capuchin monkeys, we expected the squirrel monkeys would likely be able to individuate objects by property/kind information as well as spatiotemporal information in the "magic box" task (chapter 2). Moreover, we considered both manual search and looking time measures in this chapter, the latter of which is suggested to be a more sensitive measure to reveal an individual's competence in general according to the human development literature.

However, our findings in chapter 2 are not in line with either model's prediction. The squirrel monkeys failed both spatiotemporal and property/kind individuation tasks. They neither searched or looked longer or more frequently when the outcome was unexpected. This seems to suggest squirrel monkeys do not have any form of expectation about the object(s): they were not surprised by the inconsistent quantity/number nor quality/kind of the food item(s). These findings are beyond the expectation of both the Carey-Xu and OF+PR account. Either of them will tend to expect that if the squirrel monkeys are equipped with a core knowledge system about physical objects, they would at least pass the spatiotemporal individuation task.

In view of the unexpected results that would not fit any of the existing theoretical accounts, we thought it was reasonable to have a second look before concluding that the squirrel monkeys cannot individuate objects at all, especially considering the success in capuchin monkeys, the sibling species of squirrel monkeys that shares so many aspects of its ecology. Thus, we set out to explore if the negative results may not reveal a real individuation incompetence in squirrel monkeys, but might be due to difficulties to engage in the task for intrinsic or extrinsic reasons.

Therefore, we proposed and tested two alternative possibilities, not necessarily mutually exclusive. First, we explored if the squirrel monkeys had an intrinsic tendency to react as soon as possible so that they interacted with the apparatus before thoroughly "thinking" about the event and representing the objects. Given the previous success in capuchin monkeys with the same experimental set-up, chapter 3 aimed to test what we called the "Cursory squirrel monkey vs. deliberate capuchin monkey" hypothesis by directly comparing the performance between these two species on a simple detection and hit task, the "whack-a-cricket" game. Our prediction was that, if the hypothesis was correct, then the squirrel monkeys would hit faster and/or less accurately and precisely than the capuchin monkeys. However, the results didn't support this prediction. On the contrary, the capuchins "caught" the cricket faster than the squirrel monkeys with a comparable level of accuracy and precision across the two species. The squirrel monkeys didn't act faster with a higher amount of off-target hits. Thus, the findings did not support the "Cursory squirrel monkey" hypothesis.

The second possibility is that the apparatus, set-up, and procedure in the "Magic box" experiment were less suitable for squirrel monkeys as they were initially designed for capuchin monkeys; specifically, we hypothesised that the 'magic box' would have been especially challenging for squirrel monkeys in terms of size and spatial span, the nonlinear trajectory of the objects involved, the type of events, and the overall task paradigm. Therefore, in chapter 4 we designed and conducted a set of new tasks, in which the big box was replaced by two smaller barriers placed in a smaller scenario.

A critical difference is that there were two barriers instead of one single container (the "magic box"). On the one hand, using barriers instead of a box changed the nature of this task to occlusion events rather than containment events. The latter type of event has been reported to be more difficult for both human infants and nonhuman animals including squirrel monkeys and dogs (Baillargeon & DeVos, 1991; de Blois et al., 1998; Espinosa & Buchsbaum, 2019; Hespos & Baillargeon, 2001a, 2006). On the other hand, there was either one food item or none behind each barrier, which in the OF+PR model means a resolvable qualitative disagreement between systems that should lead to successful object individuation. Hence, the two-barrier task was supposed to be a more suitable one to test squirrel monkeys' property/kind individuation ability.

Moreover, the barriers in chapter 4 could be directly knocked down by the subjects to reveal food item(s) once the apparatus was pushed forward. Thus, the event sequence was simpler than that in the "magic box" task in which there was an extra event in which the experiment removed the panel of the access hole to allow the monkey to have access to the rewards. According to Wilcox, Baillargeon, and collaborators' perspective ("event-mapping versus event monitoring" or "ongoing event versus the end of an event", we will further discuss this in section 6.1.3), this also means less task demands for the monkeys.

Finally, besides these changes from a physical reasoning point of view, the two-location paradigm of chapter 4 also returned to the original set-up used to test infants' spatiotemporal individuation ability (Spelke et al., 1995; Xu & Carey, 1996). Given that the squirrel monkeys failed the “magic box” version of the spatiotemporal individuation task in chapter 2, the two-location tasks should be an informative alternative to assess the squirrel monkeys’ spatiotemporal individuation ability.

Indeed, the findings of chapter 4 revealed that the squirrel monkeys were sensitive to both spatiotemporal and property/kind information in object individuation. When the continuity of the path (spatiotemporal information) or the kind(s) of food involved (property/kind information) suggested one object, monkeys searched preferentially behind that particular barrier first. Whereas when either information suggested two objects, one behind each barrier (i.e., in the Discontinuous condition in the spatiotemporal task, or in the Different condition in the property/kind task), monkeys showed no significant preference between the two locations.

We also found a general tendency to search both locations thoroughly: monkeys would search the other location not only when they missed the first attempt (see the “thorough search” data in chapter 4), but also when they got the food item in the first attempt (see the “second search” data in chapter 4), even in the condition where the cue indicated only one object. This pattern suggests that the squirrel monkeys tend to search redundantly, rather than hastily as we explored but failed to find evidence for in chapter 3. This redundant searching behaviour may partially explain the negative results in chapter 2 using the “magic box”, given that when there was only one region of interest for the subject to search, they would always engage in redundant searching after retrieval regardless of the conditions. In the two-location paradigm, the first search measurement crucially overcame this difficulty and uncovered evidence of individuation based on the first search choice of the monkeys. We found additional evidence that that the squirrel monkeys were more likely to search after the first retrieval when the spatiotemporal information suggested two objects, but the difference was not significant in the property/kind task. Thus, we found clear evidence that squirrel monkeys do individuate both spatiotemporally and by property/kind with our new two location task, but the results were not completely consistent across all measurements.

6.1.2 Do squirrel monkey individuate objects?

Taken together, we found three different pairs of results in squirrel monkeys depending on the tasks and measures (see Table 6-1): (1) the “magic box” tasks using the “second search” measure revealed failures in both spatiotemporal and property/kind tasks; (2) the “two-barrier” tasks using the “second search” measure revealed successful individuation with spatiotemporal but not property/kind information; (3) the “two-barrier” tasks using the “first search” measure revealed success in both spatiotemporal and property/kind tasks. To interpret these seemingly mixed results, several factors shall be considered beforehand.

First of all, as we mentioned, the two-barrier paradigm is different to the “magic box” paradigm in terms of apparatus size, trajectories of object movement, event type, and measure. The two-barrier paradigm is less demanding and more suitable for squirrel monkeys. This can (partially) explain the difference between (1) and (2) (3). The results from chapter 4 revealed that the squirrel monkeys are able to individuate objects by spatiotemporal information. This is consistent with all the existing empirical findings in other species, and also in line with the theoretical predictions. According to either the Carey-Xu account or the OF+PR model, our findings suggest that the squirrel monkeys are equipped with an object-file system, a cognitive module that emerges very early (if not innate) in human development and may have originated early in primate phylogeny.

Table 6-1 Summary of the squirrel monkeys’ individuation performance in chapter 2 and chapter 4

Task	Magic box		Two-barrier	
	Looking	Reaching	First search	Second search
Spatiotemporal OI	Fail	Fail	Pass	Pass
Property/kind OI	Fail	Fail	Pass	Fail
	(1)		(2)	(3)

Second, in relation to results (2) and (3) we noticed a ceiling effect in the “second search” measure. Two potential reasons may have caused such ceiling effect and discrepancies between these two groups of results. On the one hand, the squirrel monkeys may have a strong tendency to search redundantly for food. This sounds a reasonable strategy for squirrel monkeys, since the cost of simply searching again or around a particular place is little, especially when we deliberately optimised the apparatus to make it easier for the squirrel monkeys to search, especially in terms of the size and the way the barriers can be pushed down by the monkeys to reveal the hidden food. The tendency of redundant searching seems to be such a pervasive behaviour pattern in squirrel monkeys that this may also explain the results in chapter 2, according to which they searched generally more than capuchin monkeys did in our previous study (Kersken et al., 2020).

On the other hand, it must be noted that there were always two pieces of food in the property/kind task (two same objects versus two different objects) but not in the spatiotemporal task (one versus two objects). The change we made in the property/kind task was to rule out a possibility that the experimenter inadvertently revealed any spatiotemporal information when he manipulated two different objects. Thus, one may argue that the monkeys might have gradually learned the fact that there were always two food items, one behind each barrier, under either conditions, and therefore increased their tendency to search both locations, namely leading to an even greater ceiling in the property/kind task. However, when including the practice effect in the model, we didn't find such evidence to support this possibility. Nonetheless, a future replication including the comparison of "one object versus two different objects" will be helpful to further clarify the issue.

Last but not least, while the preferential measure we used in chapter 4 ("first search") seems to be more sensitive to reveal an individual's competence, it is an open question whether this competence is comparable to that revealed by the measures based on frequency or accumulation logic (e.g., the duration or the number of visits). For instance, it can be argued that a monkey is likely to stop searching if it has a clear and robust representation of only one piece of food, and it was already found.

The ceiling effect with the "second search" measure and the result (1) of individuation failure in the "magic box" task may support this argument. Squirrel monkeys seem to be competent when making the preferential choice ("first search"). Their object representations are strong enough to survive the challenge of conflicting information from another source, and they are able to resolve these conflicts under our experimental conditions. For example, they can individuate objects based on property/kind information ("two different objects") while overcoming the conflicting spatiotemporal information ("one object on a continuous path") in the Non-Delay-Different trials, showing no tunnel effect. However, these didn't seem to remain clear and robust representations capable of guiding their actions in the following stage of the task, leading to the ceiling effects on "second search". And they also failed the "magic box" tasks, in which the whole event and especially the coding period (or the period of time their reaction was recorded) were longer. Thus, one possibility is that the squirrel monkeys form object representations based on either source of information, but the representation may decay quickly over time.

Therefore, we presume that the squirrel monkeys' property/kind-based object representations are fragile or less robust only in terms of durability, rather than strength. This is consistent with Kibbes and Leslie's suggestion in development psychology that the featural can be bound to an object representation, but such featural binding may be more difficult to form and sustain (Kibbe & Leslie, 2011, 2019). However, our study didn't provide decisive evidence that the squirrel monkeys use spatiotemporal information better than property/kind information. It seems to be a general issue for squirrel monkeys to sustain and use over times their object representations, which is in line with the previous comparison between squirrel monkeys and capuchin monkeys using delay of gratification tasks (e.g., Anderson et al., 2010).

Hence, our current findings suggest that the squirrel monkeys are able to individuate objects based on either the spatiotemporal or the property/kind information, but their object representations may not be as durable and robust as in other primate species (including human beings, apes, macaques, and capuchin monkeys). In the next section, we try to discuss how these findings may shed light on the underlying cognitive structure(s) under two major theoretical frameworks.

6.1.3 Theoretical implications

Two major theoretical frameworks have been proposed to explain the object individuation data across various subject populations with different tasks: (1) the Carey-Xu account proposed by Xu and Carey, and (2) the OF+PR model proposed by Stavans, Baillargeon, and collaborators (Baillargeon et al., 2012; Carey & Xu, 2001; Stavans et al., 2019; Xu & Carey, 1996). Both frameworks suggest two separate systems, whereas the critical differences are about the function of the two systems ("what do these two systems each account for?") and the relationship between the two systems (stage-like or working in parallel).

First of all, they all agree that there is an object-file system that underlies spatiotemporal individuation. This system emerges early in human infants and seems to be an innate cognitive module from a core knowledge perspective (Carey & Xu, 2001; Leslie et al., 1998; Spelke, 1998; Spelke et al., 1992; Spelke & Kinzler, 2007). Thus, one could suggest that this module should have its roots in evolution and can be expected in nonhuman primates. Our findings align with such presumption. We have found that squirrel monkeys do have the ability to individuate objects by spatiotemporal information, showing an understanding of the spatiotemporal principles that a physical object follows.

The OF+PR model, in contrast to the Carey-Xu account, further suggests that the object-file system also encodes categorical information about an object (Stavans et al., 2019). However, it was beyond the scope of our current studies to explore the role of categorical information, given that all the target objects we used could be simply categorized as “food”, as well as more specific categories, such as food/fruit/raisin vs. food/nut/peanut. And it is also conceptually difficult to disentangle kind/categorical differences from featural differences (for some attempts in human infants and nonhuman primates, Cacchione et al., 2016; Cacchione et al., 2012; Phillips et al., 2010; Rakoczy & Cacchione, 2019).

In relation to property/kind individuation and its underlying mechanism, the two theoretical frameworks differ enormously in many aspects. On the one hand, Xu and Carey’s two-system account emphasized a stage-like developmental sequence in human infancy, such that younger infants would not be equipped with the second, kind-based system to individuate objects by property/kind before the age of one year. A strong version of this account even suggests that this second system relies on language acquisition (Xu, 2002, 2007). Our findings, along with the existing evidence from preverbal infants and nonhuman animal research, do not support these assumptions. The Carey-Xu account implies a time point in evolutionary history in which an individual species could individuate objects by and only by spatiotemporal information (Carey & Xu, 2001; Xu & Carey, 1996). Prior to our work such evidence (i.e., a species that succeeded spatiotemporal individuation but failed property/kind information, like a human infant between 10-month to 12-month-old) was never found in comparative research (e.g., Kersken et al., 2020; Mendes et al., 2008; Santos et al., 2002; Uller et al., 1997). Our findings in chapters 2 and 4, testing squirrel monkeys as a candidate species for such dissociation, failed to find such evidence for the Carey-Xu account. In chapter 2, squirrel monkeys failed both in spatiotemporal and property kind individuation, whereas in chapter 4 they passed both.

However, one may notice that our result (2) could be taken to match the evidence of "the gap" between the spatiotemporal and property/kind individuation (see Table 6-1). Namely, the squirrel monkeys' second search behaviour did not differ between different-object and same-object trials. However, this non-significant difference can be explained by the general high tendency to search redundantly (ceiling effect) displayed by the monkeys and the fact that they were rewarded for searching twice, since there was always a food item behind each barrier in the property/kind task but not in the spatiotemporal task. More importantly, our findings with a preferential measure (first search, or "where a monkey would search first") seem to reveal the squirrel monkey's competence of using property/kind information in object individuation. This pattern can hardly be explained under the framework of the Carey-Xu account. The lack of the second, kind-based system for object representation should lead to failure with both measures.

The OF+PR model proposed more recently by Stavans et al. (2019), on the other hand, predicts successful property/kind individuation as long as the appropriate task is applied (Stavans et al., 2019). This model suggests that the featural information is processed by the physical reasoning system as part of the event representation. Thus, successful property/kind individuation highly depends on the event(s) involved in a particular task. The tasks we used in chapter 2 ("magic box") and chapter 4 ("two-barrier") seem to conform to the task type called "the remainder task" by Stavans and collaborators (see Figure 6-1). Namely, when the event comes to an end, only one object maximum remains hidden (rather than two hidden objects in Xu and Carey's original test) behind the barrier or in the container. Thus, the disagreement between object-file and physical reasoning system is qualitative (1 vs. 0 or "nothing" vs "something") and therefore resolvable.

For example, in the "magic box" experiment, when an object A was inserted into the box and an object B was found, the object-file system would signal that there was one object in total and none left hidden according to the spatiotemporal information, whereas the physical reasoning system would encode the featural information indicating two objects in total and that one object was missing. Therefore, the disagreement between OF and PR systems can be resolved in this situation, suggesting that individuals are likely to demonstrate their ability to individuate objects by property/kind in the "magic box" task.

The rationale is similar but slightly more complex in the two-barrier task. For the last-visited location, an individual always sees the same object move around this barrier. Thus, the two systems agree that there is one and only one object in this last-visited location. The first-visited location is the more critical one in our task similar to the only hidden location in a typical "Remainder tasks": an object A moves behind the barrier and an object B moves out in the different-object condition (Stavans et al., 2019). The PR system encodes the featural distinction and indicates that one object, the object A, is left hidden behind this barrier. When the object B comes out (different-object condition), the OF system will indicate one hidden object, which is in line with the coding of the PR system. Thus, there is one object in the first-visited location as well as in the last-visited location, guiding an individual to search either location first. On the contrary, when it is an object A that comes out (same-object condition), the PR system will signal that no object is remaining hidden behind this barrier, guiding an individual to search the last-visited location first. Our findings are consistent with these expectations from the perspective of the OF+PR model.

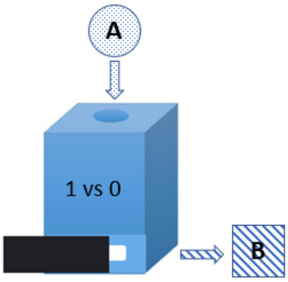
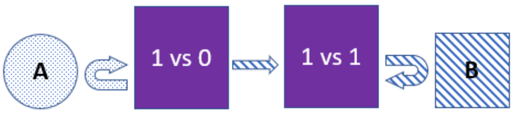
	Task (different-object condition)	OF number	PR number	OF-PR agree?	Subjects' expectation
A		0	1	Disagree qualitatively	1
B		0	1	Disagree qualitatively	1
C		Left: 0	Left: 1	Disagree qualitatively	1
		Right: 1	Right: 1	Agree	1

Figure 6-1 Subjects' performance in three types of object individuation tasks according to the OF+PR model (Stavans et al., 2019). Only the different-object condition is presented here. A is the typical "remainder task"; B is the "Magic box" task we used in chapter 2; C is the "Two-barrier task" we used in chapter 4. Note that in C, the question regarding what is hidden in the first visited location

(left location) is similar to the typical "remainder task", whereas it is beyond question that food can be found in the last visited location (right location) according to either the OF or PR system.

Moreover, an interesting result appeared when we manipulated the temporal gap when the objects were behind the first-visited location. The monkeys' preference to search the last-visited location disappeared when the object A came out after a three-second delay (Delay-Same condition). Given the lack of preference in both of the Delay conditions, one possibility is that the monkeys lose track of the whole event, leading to a random search not because of the expectation of two object items, but because they didn't have any expectation after the attention shift. As mentioned in sections 4.2.3 and 4.3.3, the possibility of an attention effect seems to be embedded in the spatiotemporal manipulation, so that it is extremely difficult to disentangle the attention effect from the expectation/representation-based behaviour. More evidence and new paradigms or techniques are required to clarify this issue.

However, there is an alternative explanation worth noting: presumably, the OF system may have indicated that two different objects of the same kind were involved in the event, while one of them remained hidden behind the barrier. Thus, the OF system (signalling one remaining) and the PR system (signalling none remaining) would disagree qualitatively, leading to the current result that monkeys search the first-visited location first as often as the last-visited location. This would be an interesting finding, if this explanation is correct, since this is probably the first case in which the OF system indicates more objects and overrides the PR system. Note that in most (if not all) of the existing experimental scenarios, the number of object(s) suggested by the property/kind information is no fewer than that suggested by the spatiotemporal-based object-file system.

We found that the current result can be well interpreted with the object-file system which both the Carey-Xu account and the OF+PR model include: a new object-file is opened and can only be opened based on spatiotemporal information (at the moment of the delay), but not the property/kind information (Carey & Xu, 2001; Stavans et al., 2019). However, we also noticed a contrary effect: in the Non-Delay-Different-Object condition, when the spatiotemporal information indicated one object while the property/kind information indicated two. These results involving conflicting information cannot fully be explained by the Carey-Xu account, in which the spatiotemporal information plays a predominant role (Carey & Xu, 2001; Flombaum et al., 2012). In contrast, these results may be better interpreted under the framework of the OF+PR model, and even supplement this model by adding a new and informative kind of trial. Note that the OF+PR model emphasizes resolving the conflict between systems without clarifying how (Stavans et al., 2019). Currently there can be two possible interpretations. That is, to resolve a disagreement (if it is resolvable), an individual may either adopt the output of the PR system or to choose the larger value between the two systems. In most of the previous studies, these two answers were entangling with each other, when the output of the PR system is also the larger value (e.g., in a different-object condition of standard Xu-Carey individuation task, the property/kind information indicates more objects than the spatiotemporal system does). Our results here seem to support the latter, i.e., an individual tend to adopt the larger value provided by the two systems, implying that both spatiotemporal and property/kind information contribute to object individuation.

Taken together, our findings in squirrel monkeys favour the OF+PR model more than the Carey-Xu account, at least when describing the representations of squirrel monkeys. The OF+PR model may better explain our current findings in detail, especially that the squirrel monkeys individuate objects by property/kind. Meanwhile, the stage-like dissociation between spatiotemporal and property/kind individuation is not found in our study, which again casts doubt on the Carey-Xu account. Nonetheless, we of course agree that the absence of evidence is not the same as evidence of absence. It is, therefore, possible that a dissociation between spatiotemporal and property/kind individuation, similar to the developmental gap originally reported in human infants, can be found in a different species of primates or other animal species. In the meantime, our results suggest that both types of individuation are present at the same time in primates at least from as early as the common ancestor of squirrel monkeys and humans.

6.2 Capuchin monkeys represent a hidden causal agent

The studies presented in chapters 2 to 4 investigated whether squirrel monkeys can track and represent hidden objects based on either spatiotemporal information or property/kind information. Our final experiment explored a different, but related ability in capuchin monkeys, namely, if they can track and represent hidden agents based on causal information, i.e., the agent's causal capability to initiate events as reflected in the observable physical events he caused.

6.2.1 Empirical findings

Our findings in chapter 5 support the idea that capuchin monkeys possess the cognitive ability to infer agents from the consequences of their actions. The capuchin monkeys were more likely to choose the experimenter's side in the agentive trials, when the hidden experimenter directly initiated an event seemingly from behind one of the screens, compared to the arbitrary control trials, when the event was caused by a shake of the table. Thus, in our experimental context, the observed events served as the cues, which were either effective (agentive condition) or ambiguous (arbitrary control condition) to indicate the location of the causal agent. The different behavioural response between these two conditions suggested that the capuchin monkeys encoded these cues and reasoned about the cause of the event accordingly. When the cause was a shake of the table, the location of the experimenter was irrelevant. The structure of the physical event was complete: the shake caused the event in which the stimulus rolled down the ramp or fell, so, in this case, they needed not to and could not infer the hiding location of the experimenter. Thus, in this case the monkeys tended to track the stimulus and went to search the box on that direction, showing a movement endpoint bias. In contrast, when the cause was hidden from the monkeys but there were causal cues about its location, the capuchin monkeys tended to posit an unobserved causal source to account for the event, like human infants in the developmental literatures (Muentener & Carey, 2010; Muentener & Schulz, 2014; Wu et al., 2016). As a consequence, they were able to trace back to the source to search the location where the experimenter should be, overcoming the endpoint bias.

However, this interpretation has a soft spot: the monkeys' performance on the Agentive trials didn't significantly exceed the chance level. However, the fact that performance on the Arbitrary control trials was significantly below chance needs to be accounted for. Our suggested interpretation is that the monkeys in general tend to track the stimulus to the "wrong" side. Two features of the movement events were not perfectly controlled in our experiment and either of them could contribute to the bias in the current results. For example, the direction of the movement may be better manipulated in future studies. Among the four pairs of movement types in the current study, there was only the “raking” pair in which the stimulus moved *towards* the correct location (where the agent should be). A post-hoc analysis found that the monkeys’ performance in these Agentive trials was above chance (marginally significant). We presume that with a larger number of trials and larger sample size, a future study using this paradigm may found more significant results. Second, the destination or endpoint of the stimulus may be better manipulated in the future studies. Among the four pairs of movement types in the current study, there was only the “rolling” pair in which the stimulus moved *out* of the stage disappearing behind the screen. We assume that when the stimulus left the stage and became invisible to the monkeys, it would attract more attention, leading to a stronger bias towards the side at which the stimulus had disappeared. This assumption was supported by another post-hoc analysis focusing only on the rolling conditions. In summary, we suggest that by better manipulating these two features of movement future studies may be able to produce a clearer result, in which the agent-related inference can be dissociated from the endpoint bias.

The endpoint bias was not only reported in the current study. At least two lines of developmental research have emphasized the difference between tracing back to the causal source and tracking towards the outcome or goal of an event. On the one hand, Csibra and Gergely termed the ability to predict the outcome or goal of an action as “teleological inference”, to differentiate it from the ability of causal explanation or the ability to reason about a mentalistic cause (Csibra & Gergely, 1998; Gergely & Csibra, 2003). They further suggested that the former ability is the prerequisite ability or origin of the latter. Accordingly, it is possible that inferring the cause based on the effect, compared to predicting the effect based on the cause, may reflect a higher level of causal understanding. On the other hand, Lakusta and collaborators found that human infants showed a general tendency to attend to the endpoint or the goal, compared to the starting-point or the source of a motion event (Lakusta & Carey, 2015; Lakusta & DiFabrizio, 2017; Lakusta, Muentener, Petrillo, Mullanaphy, & Muniz, 2017). This “endpoint/goal bias” seemed to be common in both preverbal and verbal human children, and seems to coincide with that shown in our findings with the capuchin monkeys in the control conditions. This bias or attentional asymmetry could be modulated by the agency and causal structure of the event. For example, in our study the endpoint bias was weakened when the starting point caused the event (e.g., a spring in a box that launched a ball), suggesting an interest on the cause (Lakusta et al., 2017).

Furthermore, in the developmental literature the goal bias is stronger when the moving target is agentive compared to when the target is non-agentive (Lakusta & DiFabrizio, 2017; Lakusta et al., 2017). That is, infants attended the starting point more when the moving figure was an agent (e.g., a duck) than when it was a physical object (e.g., a leaf). One possible interpretation is that the infants can and tend to seek a causal explanation of the events especially when the cause of the event is unclear, i.e., when the target is inert, which is in line with the findings of Saxe and colleagues (Saxe et al., 2005, 2007). In other words, the infants posit the missing part, the cause or causal source, of an event. Our current findings in chapter 5 suggest that the capuchin monkeys have a similar ability to understand causality and represent causal event in a gestalt way.

Our finding may suggest that capuchin monkeys are able to locate a hidden causal agent based on an observed causal event (a strong version of interpretation), or that capuchin monkeys are sensitive to the causal event caused by an agent (a weaker version of interpretation). This could imply a set of homologous cognitive module(s) shared with human infants. Such module(s) enable an individual to represent an agent who is not directly perceived based on a "top-down expectation-driven inference of the best explanation", rather than based on perceptual features in a bottom-up way (Saxe et al., 2005). Namely, in our experiment, the causal role is posited and attributed to the hidden agent, along with the enduring causal property to start an event, so that one can complete the representation of this causal event in a Gestalt manner. From this point of view, it is reasonable to situate agent representations within a broader framework of event representations (or more specifically, the representation of goal-directed action, as suggested by Csibra and Gergely and agreed by the Core Knowledge Systems account) (Csibra & Gergely, 1998; Gergely & Csibra, 2003; Spelke & Kinzler, 2007).

This would be analogous to how object representations are related to the physical reasoning system, according to Baillargeon and colleagues (Baillargeon et al., 2012; Stavans et al., 2019). In the latter case, the physical reasoning system encodes the property information of the object to decide whether a property change is possible or impossible during certain types of event (Baillargeon et al., 2012; Stavans et al., 2019; Wang & Baillargeon, 2008). When doing so, the physical reasoning system contributes to the ability of object individuation. For agent representation, the causal role and the causal property (or causal information) of an agent may also be encoded by a similar (if not the same) reasoning system under the event-dependent context. An advantage of this interpretation is that it provides an integrative perspective on agents and related action/events, whereas the present literature tends to treat agents and actions either separately or in a confounded way. For instance, as Saxe and colleagues pointed out, Gergely and Csibra underscore so much the goal-directed action that they contend the goal is attributed to the action itself rather than to the agent (Gergely & Csibra, 1999; Saxe et al., 2007). Another example is how, in the core knowledge account, Spelke and colleagues proposed a core representational system of "agents and their actions", without further differentiation (Spelke & Kinzler, 2007).

6.2.2 Limitations and future directions

First of all, as we mentioned in section 5.4 and section 6.2.1, to verify our current finding, follow-up experiment(s) would be required. At least two features of the movement need to be manipulated or controlled in order to disentangle the endpoint bias and the reasoning about a hidden causal agent. Hereby we would like to suggest a rough example of how this may be done. On the one hand, a future study may expand the number of movement types, matching the number of conditions in which the stimulus moves *from* and *towards* the agent's side. Or one may use *raking* versus *pushing* conditions, but increasing the number of trials under each condition. On the other hand, to reduce or eliminate the endpoint bias, the stimulus could be stopped somewhere on the stage, thereby staying visible to the subject. Conversely, one could increase the set of conditions in which the stimulus moves into hiding, and to compare with the on-stage conditions, in order to explore the effect of the endpoint bias.

Another future direction is to test the monkeys' ability in a more direct way. It is difficult to directly test nonhuman animals' inferential abilities about non-food entities, in this case a human agent, since food rewards are usually the main source of motivation for an animal to participate in an experiment. To cope with this challenge, we tried to establish an association between the agent and the food, but the capuchin monkeys' performance that we measured was their search for the food rewards. Thus, it could be argued that in chapter 5 we only assessed the agent inference by the monkeys in an indirect way, with the food serving as the intermediate link. However, first of all, we argue that the search preference did reflect the subjects' inference of the agent's presence. Note that the capuchin monkeys were used to get food rewards from the experimenters before this study, and they also learned (or showed their existing tendency) to approach this particular experimenter for food in the current study during the familiarization phase. Furthermore, the actual location of the food rewards or other cues related to the food (such as odours or sounds) were not sufficient to explain the different searching patterns across conditions, given that the food was hidden beforehand on the side where the experimenter should be in the agentive trials, or on the equivalent side in the control trials.

Nonetheless, we suggest that future studies can try to measure the inference of agents more directly. Two potential options have been illustrated in the literature. On the one hand, Saxe and colleagues used the violation of expectation paradigm in their infant studies. With this looking time measure, whether or not there is an agent hidden (or whether it is an agent or an object) is revealed to the subjects. Thus, the reaction (looking time) of the subject can be interpreted as a direct evidence of what a subject is predicting and whether it is an agent. Our studies in chapter 2 along with some previous literature suggest that applying such techniques in nonhuman primates is feasible (e.g., Kersken et al., 2019; Krupenye et al., 2016).

Thus, instead of the current version of tracing a food-provider, our task could be adapted to a violation of expectation paradigm, where additional control conditions could be incorporated. For example, after watching a causal event, the subject either sees an agent or an inert object appearing in the critical location. In line with our current findings, we would expect a capuchin monkey to be surprised and look longer if an inert object appears instead of an agent.

On the other hand, Taylor and colleagues used a threatening context in their crow study, in which the agent was a threat to the foraging subject instead of a "helper" that distributed the food rewards as we did in our experiment. Thus, it could be argued that the crows' reaction towards the agent was directly assessed in their experiment. In addition, compared to the helping-reinforcing context we used here, a threatening or competing context has been suggested to be more suitable to test nonhuman primate species in some literature involving predicting agents' behaviours (for an example in capuchin study, Hare et al., 2003), despite the concerns that threatening or competing contexts may stress the subjects, as shown in our study.

In our experiment the capuchin monkeys had to infer the *location* rather than the *existence* of the hidden agent. Therefore, one needs to be cautious when comparing the capuchin monkeys' competence with the human infants. Our findings show that the monkeys can attribute the cause of an observed event to an agent that they already knew was present, but not where exactly. We do not know if they could infer the existence of a completely unknown agent from the causal information available in observed events, as human infants do (Saxe et al., 2005, 2007).

For practical reasons, it was not possible for us to implement this version of the paradigm. First, the experimenter had to introduce the monkey into the cubicle for testing. Second, it would have been potentially frightening for at least some of the monkeys if no experimenter was present while they were in the research cubicle, separated from the group. Indeed, one subject (Flojo) quitted the experiment because he was frightened even when he watched the experimenter hiding behind the screen.

In any case, despite the monkeys benefitting from the previous knowledge of the presence of the experimenter, our paradigm still required that they inferred his exact location on the basis of understanding the properties of agentive causality, overcoming other biases and superficial information. Our findings suggest that capuchin monkeys possess such an understanding.

6.3 Conclusion

In sum, our series of studies explored the ability to represent and track objects and agents in two new world monkey species, the squirrel monkey and the capuchin monkey. Our results suggest that squirrel monkeys can individuate objects based on both spatiotemporal and property/kind information when tested with an appropriate paradigm, adding to the body of literature suggesting that both types of individuation go hand in hand in evolution and development and are present in all primate species tested so far. Importantly, we showed that the type of paradigm used to test for this ability matters, and that apparent failures to individuate may be due to artefactual limitations in what different species find challenging in testing.

We also demonstrated for the first time that capuchin monkeys can infer the location of a hidden agent based on the causal structure of the information contained in an event. Taken together, our findings suggest that these core knowledge abilities for object and agent representation are already present in these distant primate relatives of human beings, and therefore have a deep root in the evolutionary history of primates.

Reference

Abdai, J., Ferdinandy, B., Terencio, C. B., Pogány, Á., & Miklósi, Á. (2017). Perception of animacy in dogs and humans. *Biology Letters*, *13*(6). https://doi.org/10.1098/rsbl.2017.0156

Addessi, E., Paglieri, F., Beran, M. J., Evans, T. A., Macchitella, L., De Petrillo, F., & Focaroli, V. (2013). Delay choice versus delay maintenance: Different measures of delayed gratification in capuchin monkeys (cebus apella). *Journal of Comparative Psychology*, *127*(4), 392–398. https://doi.org/10.1037/a0031869

Addessi, E., Paglieri, F., & Focaroli, V. (2011). The ecological rationality of delay tolerance: Insights from capuchin monkeys. *Cognition*, *119*(1), 142–147. https://doi.org/10.1016/j.cognition.2010.10.021

Aguiar, A., & Baillargeon, R. (1999). 2.5-Month-Old Infants' Reasoning about When Objects Should and Should Not Be Occluded. *Cognitive Psychology*, *39*(2), 116–157. https://doi.org/10.1006/cogp.1999.0717

Amici, F., Cacchione, T., & Bueno-Guerra, N. (2017). Understanding of object properties by sloth bears, Melursus ursinus ursinus. *Animal Behaviour*, *134*, 217–222. https://doi.org/10.1016/J.ANBEHAV.2017.10.028

Anderson, J. R., Kuroshima, H., & Fujita, K. (2010). Delay of Gratification in Capuchin Monkeys (Cebus apella) and Squirrel Monkeys (Saimiri sciureus). *Journal of Comparative Psychology*, *124*(2), 205–210. https://doi.org/10.1037/a0018240

Baillargeon, R. (1995). Physical reasoning in infancy. In *The Cognitive Neurosciences* (Vol. 9, pp. 181–204). Retrieved from http://internal.psychology.illinois.edu/infantlab/articles/Baillargeon_1995b.pdf

Baillargeon, R., & Carey, S. (2012). Core cognition and beyond: The acquisition of physical and numerical knowledge. In *Early childhood development and later outcome.* (pp. 33–65). Retrieved from https://www.researchgate.net/profile/Renee_Baillargeon/publication/266504739_Core_Cognition_and_Beyond_The_Acquisition_of_Physical_and_Numerical_Knowledge/links/545ab6e10cf2c46f6643877e.pdf

Baillargeon, R., & DeVos, J. (1991). Object Permanence in Young Infants: Further Evidence. *Child Development*, *62*(6), 1227–1246. https://doi.org/10.1111/j.1467-8624.1991.tb01602.x

Baillargeon, R., Stavans, M., Wu, D., Gertner, Y., Setoh, P., Kittredge, A. K., & Bernard, A.

(2012). Object Individuation and Physical Reasoning in Infancy: An Integrative Account. *Language Learning and Development*, *8*(1), 4–46. https://doi.org/10.1080/15475441.2012.630610

Baillargeon, R., Wu, D., Yuan, S., Li, J., & Luo, Y. (2009). Young infants' expectations about self-propelled objects. In *The Origins of Object Knowledge* (pp. 285–352). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199216895.003.0012

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329–349. https://doi.org/10.1016/j.cognition.2009.07.005

Barnes, J. L., Hill, T., Langer, M., Martinez, M., & Santos, L. R. (2008). Helping behaviour and regard for others in capuchin monkeys (Cebus apella). *Biol. Lett*, *4*, 638–640. https://doi.org/10.1098/rsbl.2008.0410

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling Versus Unable: Infants' Understanding of Intentional Action. https://doi.org/10.1037/0012-1649.41.2.328

Berthier, N. E., Bertenthal, B. I., Seaks, J. D., Sylvia, M. R., Johnson, R. L., & Clifton, R. K. (2001). Using Object Knowledge in Visual Tracking and Reaching. *Infancy*, *2*(2), 257–284. https://doi.org/10.1207/S15327078IN0202_9

Biro, S., & Leslie, A. M. (2007). Infants' perception of goal-directed actions: Development through cue-based bootstrapping. *Developmental Science*, *10*(3), 379–398. https://doi.org/10.1111/j.1467-7687.2006.00544.x

Bonatti, L., Frot, E., & Mehler, J. (2005). What face inversion does to infants' counting abilities. *Psychological Science*, *16*(7), 506–510. https://doi.org/10.1111/j.0956-7976.2005.01565.x

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

Boogert, N. J., Arbilly, M., Muth, F., & Seed, A. M. (2013). Do crows reason about causes or agents? The devil is in the controls. *Proceedings of the National Academy of Sciences*, *110*(4), E273–E273. https://doi.org/10.1073/pnas.1219664110

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

Burke, L. (1952). On the Tunnel Effect. *Quarterly Journal of Experimental Psychology*, *4*(3), 121–138. https://doi.org/10.1080/17470215208416611

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., 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

Call, J. (2004). Inferences about the location of food in the great apes (Pan paniscus, Pan troglodytes, Gorilla gorilla, and Pongo pygmaeus). *Journal of Comparative Psychology*, *118*(2), 232–241. https://doi.org/10.1037/0735-7036.118.2.232

Call, J. (2007). Apes know that hidden objects can affect the orientation of other objects. *Cognition*, *105*(1), 1–25. https://doi.org/10.1016/j.cognition.2006.08.004

Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). “Unwilling” versus “unable”: Chimpanzees’ understanding of human intentional action. *Developmental Science*, *7*(4), 488–498. https://doi.org/10.1111/j.1467-7687.2004.00368.x

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2008.02.010

Canteloup, C., & Meunier, H. (2017). ‘Unwilling’ *versus* ‘unable’: Tonkean macaques’ understanding of human goal-directed actions. *PeerJ*, *5*, e3227. https://doi.org/10.7717/peerj.3227

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

Civelek, Z., Call, J., & Seed, A. M. (2020). Inferring Unseen Causes: Developmental and Evolutionary Origins. *Frontiers in Psychology*, *11*, 872. https://doi.org/10.3389/fpsyg.2020.00872

Crichton, M. T., & Lange-Küttner, C. (1999). Animacy and propulsion in infancy: Tracking, waving and reaching to self-propelled and induced moving objects. *Developmental Science*, *2*(3), 318–324. https://doi.org/10.1111/1467-7687.00077

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, *107*(2), 705–717. https://doi.org/10.1016/j.cognition.2007.08.001

Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science*, *1*(2), 255–259. https://doi.org/10.1111/1467-7687.00039 de Blois, S. T., Novak, M. A., & Bond, M. (1998). Object permanence in orangutans (Pongo pygmaeus) and squirrel monkeys (Saimiri sciureus). *Journal of Comparative Psychology (Washington, D.C. : 1983)*, *112*(2), 137–152. https://doi.org/10.1037//0735-7036.112.2.137

Di Giorgio, E., Lunghi, M., Vallortigara, G., & Simion, F. (2021). Newborns' sensitivity to speed changes as a building block for animacy perception. *Scientific Reports*, *11*(1), 1–10. https://doi.org/10.1038/s41598-020-79451-3

Drayton, L. A., & Santos, L. R. (2014). Capuchins' (Cebus apella) sensitivity to others' goal-directed actions in a helping context. *Animal Cognition*, *17*(3), 689–700. https://doi.org/10.1007/s10071-013-0700-5

Dymond, S., Haselgrove, M., & McGregor, A. (2013). Clever crows or unbalanced birds? *Proceedings of the National Academy of Sciences of the United States of America*, *110*(5), E336. https://doi.org/10.1073/pnas.1218931110

Espinosa, J., & Buchsbaum, D. (2019). Domestic dog understanding of containment and occlusion events. In *Cognitive Science* (p. 3451).

Evans, T. A., Beran, M. J., Paglieri, F., & Addessi, E. (2012). Delaying gratification for food and tokens in capuchin monkeys (Cebus apella) and chimpanzees (Pan troglodytes): When quantity is salient, symbolic stimuli do not improve performance. *Animal Cognition*, *15*(4), 539–548. https://doi.org/10.1007/s10071-012-0482-1

Flombaum, J. I., Kundey, S. M., Santos, L. R., & Scholl, B. J. (2004). Dynamic object individuation in rhesus macaques: A study of the tunnel effect. *Psychological Science*, *15*(12), 795–800. https://doi.org/10.1111/j.0956-7976.2004.00758.x

Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 840–853. https://doi.org/10.1037/0096-1523.32.4.840

Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2012). Spatiotemporal priority as a fundamental principle of object persistence. In *The Origins of Object Knowledge*. https://doi.org/10.1093/acprof:oso/9780199216895.003.0006

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

Fragaszy, D. M. (1985). Cognition in Squirrel Monkeys A Contemporary Perspective. In *Handbook of Squirrel Monkey Research* (pp. 55–98). Springer US. https://doi.org/10.1007/978-1-4757-0812-7_3

Fragaszy, D. M., Visalberghi, E., & Fedigan, L. M. (2004). *The complete capuchin: the biology of the genus Cebus.* Cambridge: Cambridge University Press.

French, G. M. (1959). Performance of squirrel monkeys on variants of delayed response. *Journal of Comparative and Physiological Psychology*, *52*(6), 741–745. https://doi.org/10.1037/h0038973

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11), 1325–1330. https://doi.org/10.1111/2041-210X.12584

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

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: the naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292. https://doi.org/10.1016/S1364-6613(03)00128-1

Gómez, J.-C. (2004). *Apes, monkeys, children and the growth of mind.* Harvard University Press. Retrieved from http://psycnet.apa.org/psycinfo/2004-18822-000

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

Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Current Biology*. https://doi.org/10.1016/j.cub.2008.02.039

Hare, B., Addessi, E., Call, J., Tomasello, M., & Visalberghi, E. (2003). Do capuchin monkeys, Cebus apella, know what conspecifics do and do not see? *Animal Behaviour*, *65*(1), 131–142. https://doi.org/10.1006/ANBE.2002.2017

Haun, D. B. M., Jordan, F. M., Vallortigara, G., & Clayton, N. S. (2010). Origins of spatial, temporal and numerical cognition: Insights from comparative psychology. *Trends in Cognitive Sciences*, *14*, 552–560. https://doi.org/10.1016/j.tics.2010.09.006

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

Hauser, M. D., & Spelke, E. S. (2004). Evolutionary and developmental foundations of human knowledge I . The core knowledge thesis II . Comparative evolutionary approaches. *The Cognitive Neurosciences*, *3*(1523), 853–864. https://doi.org/10.1098/rspb.2003.2414

Hespos, S. J., & Baillargeon, R. (2001a). Infants' Knowledge about Occlusion and Containment Events: A Surprising Discrepancy. *Psychological Science*, *12*(2), 141–147. https://doi.org/10.1111/1467-9280.00324

Hespos, S. J., & Baillargeon, R. (2001b). Reasoning about containment events in very young infants. *Cognition*, *78*(3), 207–245. https://doi.org/10.1016/S0010-0277(00)00118-9

Hespos, S. J., & Baillargeon, R. (2006). Décalage in infants' knowledge about occlusion and containment events: Converging evidence from action tasks. *Cognition*, *99*(2). https://doi.org/10.1016/j.cognition.2005.01.010

Hill, A., Collier-Baker, E., & Suddendorf, T. (2012). Inferential reasoning by exclusion in children (Homo sapiens). *Journal of Comparative Psychology*, *126*(3), 243–254. https://doi.org/10.1037/a0024449

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

Jaakkola, K. (2014). Do Animals Understand Invisible Displacement? A Critical Review. *Journal of Comparative Psychology*, *128*(2), 1–15. https://doi.org/10.1037/a0035675

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

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Johnson, S. C. (2003). Detecting agents. *Philosophical Transactions of the Royal Society B:*

Biological Sciences, *358*(1431), 549–559. https://doi.org/10.1098/rstb.2002.1237

Johnson, S. C., Shimizu, Y. A., & Ok, S. J. (2007). Actors and actions: The role of agent behavior in infants' attribution of goals. *Cognitive Development*, *22*(3), 310–322. https://doi.org/10.1016/j.cogdev.2007.01.002

Johnson, S. C., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze following in 12-month-olds. *Developmental Science*, *1*(2), 233–238. https://doi.org/10.1111/1467-7687.00036

Kamewari, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. (2005). Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. *Cognitive Development*, *20*(2), 303–320. https://doi.org/10.1016/j.cogdev.2005.04.004

Kersken, V., Zhang, D., Gomez, J.-C., Seed, A. M., & Ball, D. (2020). Capuchin Monkeys Individuate Objects Based on Spatio-temporal and Property/Kind Information: Evidence from Looking and Reaching Measures. *Animal Behavior and Cognition*, *7*(3), 343–364. https://doi.org/10.26451/abc.07.03.06.2020

Kibbe, M. M., & Leslie, A. M. (2011). What do infants remember when they forget? location and identity in 6-month-olds' memory for objects. *Psychological Science*, *22*(12), 1500–1505. https://doi.org/10.1177/0956797611420165

Kibbe, M. M., & Leslie, A. M. (2019). Conceptually Rich, Perceptually Sparse: Object Representations in 6-Month-Old Infants' Working Memory. *Psychological Science*, *30*(3), 362–375. https://doi.org/10.1177/0956797618817754

Krøjgaard, P. (2000). Object individuation in 10-month-old infants: Do significant objects make a difference? *Cognitive Development*, *15*(2), 169–184. https://doi.org/10.1016/s0885-2014(00)00020-4

Krøjgaard, P. (2003). Object individuation in 10-month-old infants: Manipulating the amount of introduction. *British Journal of Developmental Psychology*, *21*(3), 447–463. https://doi.org/10.1348/026151003322277793

Krøjgaard, P. (2007). Comparing infants' use of featural and spatiotemporal information in an object individuation task using a new event-monitoring design. *Developmental Science*, *10*(6), 892–909. https://doi.org/10.1111/j.1467-7687.2007.00640.x

Krupenye, C., & Call, J. (2019). Theory of mind in animals: Current and future directions. *Wiley Interdisciplinary Reviews: Cognitive Science*, *10*(6), 1–25. https://doi.org/10.1002/wcs.1503

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, *354*(6308), 110–114. https://doi.org/10.1126/science.aaf8110

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2017). A test of the submentalizing hypothesis: Apes' performance in a false belief task inanimate control. *Communicative & Integrative Biology*, *10*(4), e1343771. https://doi.org/10.1080/19420889.2017.1343771

Kuhlmeier, V. A., Bloom, P., & Wynn, K. (2004). Do 5-month-old infants see humans as material objects? *Cognition*, *94*(1), 95–103. https://doi.org/10.1016/j.cognition.2004.02.007

Lakusta, L., & Carey, S. (2015). Twelve-Month-Old Infants' Encoding of Goal and Source Paths in Agentive and Non-Agentive Motion Events. *Language Learning and Development*, *11*(2), 152–175. https://doi.org/10.1080/15475441.2014.896168

Lakusta, L., & DiFabrizio, S. (2017). And, the Winner Is…A Visual Preference for Endpoints over Starting Points in Infants' Motion Event Representations. *Infancy*, *22*(3), 323–343. https://doi.org/10.1111/infa.12153

Lakusta, L., Muentener, P., Petrillo, L., Mullanaphy, N., & Muniz, L. (2017). Does Making Something Move Matter? Representations of Goals and Sources in Motion Events With Causal Sources. *Cognitive Science*, *41*(3), 814–826. https://doi.org/10.1111/cogs.12376

Leonardi, R., Buchanan-Smith, H. M., Dufour, V., MacDonald, C., & Whiten, A. (2010). Living together: Behavior and welfare in single and mixed species groups of capuchin (Cebus apella) and squirrel monkeys (Saimiri sciureus). *American Journal of Primatology*, *72*(1), 33–47. https://doi.org/10.1002/ajp.20748

Leslie, A. M. (1996). A theory of agency. In *Causal Cognition* (pp. 121–141). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198524021.003.0005

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*(3), 265–288. https://doi.org/10.1016/S0010-0277(87)80006-9

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

Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science*, *358*(6366), 1038–1041. https://doi.org/10.1126/science.aag2132

Luncz, L. V., de la Torre, I., Proffitt, T., Falótico, T., Haslam, M., & Ottoni, E. B. (2016).

Wild monkeys flake stone tools. *Nature*, *539*(7627), 85–88. https://doi.org/10.1038/nature20112

Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Developmental Science*, *14*(2), 453–460. https://doi.org/10.1111/j.1467-7687.2010.00995.x

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? - Psychological reasoning in 5-month-old infants. *Psychological Science*, *16*(8), 601–608. https://doi.org/10.1111/j.1467-9280.2005.01582.x

Luo, Y., Kaufman, L., & Baillargeon, R. (2009). Young infants’ reasoning about physical events involving inert and self-propelled objects. *Cognitive Psychology*, *58*(4), 441–486. https://doi.org/10.1016/j.cogpsych.2008.11.001

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

MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., … Burkart, J. M. (2014). The evolution of self-control. https://doi.org/10.1073/pnas.1323533111

Mash, C., Novak, E., Berthier, N. E., & Keen, R. (2006). What do two-year-olds understand about hidden-object events? *Developmental Psychology*, *42*(2), 263–271. https://doi.org/10.1037/0012-1649.42.2.263

Mathieu, M., Bouchard, M. A., Granger, L., & Herscovitch, J. (1976). Piagetian object-permanence in Cebus capucinus, Lagothrica flavicauda and Pan troglodytes. *Animal Behaviour*, *24*(3), 585–588. https://doi.org/10.1016/S0003-3472(76)80071-1

McCurry, S., Wilcox, T., & Woods, R. (2009). Beyond the search barrier: A new task for assessing object individuation in young infants. *Infant Behavior and Development*, *32*(4), 429–436. https://doi.org/10.1016/j.infbeh.2009.07.002

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

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia*, *46*(3), 774–785. https://doi.org/10.1016/j.neuropsychologia.2007.10.005

Moura, A. C. de A., & Lee, P. C. (2004). Capuchin stone tool use in Caatinga dry forest. *Science (New York, N.Y.)*, *306*(5703), 1909. https://doi.org/10.1126/science.1102558

Muentener, P., Bonawitz, E., Horowitz, A., & Schulz, L. (2012). Mind the gap: Investigating toddlers’ sensitivity to contact relations in predictive events. *PLoS ONE*, *7*(4), e34061.

https://doi.org/10.1371/journal.pone.0034061

Muentener, P., & Carey, S. (2010). Infants' causal representations of state change events. *Cognitive Psychology*, *61*(2), 63–86. https://doi.org/10.1016/j.cogpsych.2010.02.001

Muentener, P., & Schulz, L. (2014). Toddlers infer unobserved causes for spontaneous events. *Frontiers in Psychology*, *5*(DEC), 1496. https://doi.org/10.3389/fpsyg.2014.01496

Munakata, Y. (2001, July 1). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*. Elsevier Current Trends. https://doi.org/10.1016/S1364-6613(00)01682-X

Munakata, Y., Santos, L. R., Spelke, E. S., Hauser, M. D., & O'Reilly, R. C. (2001). Visual Representation in the Wild: How Rhesus Monkeys Parse Objects. *Journal of Cognitive Neuroscience*, *13*(1), 44–58. https://doi.org/10.1162/089892901564162

Needham, A., & Baillargeon, R. (2000). Infants' use of featural and experiential information in segregating and individuating objects: A reply to Xu, Carey and Welch (2000). *Cognition*. https://doi.org/10.1016/S0010-0277(99)00075-X

Neilands, P. D., Jelbert, S. A., Breen, A. J., Schiestl, M., Taylor, A. H., & Gopnik, A. (2016). How Insightful Is 'Insight'? New Caledonian Crows Do Not Attend to Object Weight during Spontaneous Stone Dropping. *PLOS ONE*, *11*(12), e0167419. https://doi.org/10.1371/journal.pone.0167419

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

Phillips, W., Barnes, J. L., Mahajan, N., Yamaguchi, M., & Santos, L. R. (2009). "Unwilling" versus "unable": Capuchin monkeys' (Cebus apella) understanding of human intentional action. *Developmental Science*, *12*(6), 938–945. https://doi.org/10.1111/j.1467-7687.2009.00840.x

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

Poulin-Dubois, D., Lepage, A., & Ferland, D. (1996). Infants' concept of animacy. *Cognitive Development*, *11*(1), 19–36. https://doi.org/10.1016/S0885-2014(96)90026-X

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*(1–2), 127–158. https://doi.org/10.1016/S0010-0277(00)00156-6

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

Rhemtulla, M., & Xu, F. (2007). Sortal concepts and causal continuity: comment of Rips, Blok, and Newman (2006). *Psychological Review*, *114*(4), 1087–1094. https://doi.org/10.1037/0033-295X.114.4.1087

Rivera, S. M., & Zawaydeh, A. N. (2007). Word comprehension facilitates object individuation in 10- and 11-month-old infants. *Brain Research*, *1146*(1), 146–157. https://doi.org/10.1016/j.brainres.2006.08.112

Rosa-Salva, O., Grassi, M., Lorenzi, E., Regolin, L., & Vallortigara, G. (2016). Spontaneous preference for visual cues of animacy in naïve domestic chicks: The case of speed changes. *Cognition*, *157*, 49–60. https://doi.org/10.1016/j.cognition.2016.08.014

Roth, G., & Dicke, U. (2005). Evolution of the brain and intelligence. *Trends in Cognitive Sciences*, *9*(5), 250–257. https://doi.org/10.1016/j.tics.2005.03.005

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., & Hood, B. M. (2012). Object representation as a central issue in cognitive science. In *The Origins of Object Knowledge* (pp. 1–24). https://doi.org/10.1093/acprof:oso/9780199216895.003.0001

Santos, L. R., Seelig, D., & Hauser, M. D. (2006). Cotton-top tamarins' (Saguinus oedipus) expectations about occluded objects: A dissociation between looking and reaching tasks. *Infancy*, *9*(2), 147–171. https://doi.org/10.1207/s15327078in0902_4

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

Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science*, *16*(12), 995–1001. https://doi.org/10.1111/j.1467-9280.2005.01649.x

Saxe, R., Tzelnic, T., & Carey, S. (2007). Knowing Who Dunnit: Infants Identify the Causal Agent in an Unseen Causal Interaction. *Developmental Psychology*, *43*(1), 149–158.

https://doi.org/10.1037/0012-1649.43.1.149

Schino, G., Spinozzi, G., & Berlinguer, L. (1990). *Object Concept and Mental Representation in Cebus apella and Macaca fascicularis*. *PRIMATES* (Vol. 31).

Schneider, G. E. (1969). Two visual systems. *Science*, *163*(3870), 895–902. https://doi.org/10.1126/science.163.3870.895

Scholl, B. J., & Pylyshyn, Z. W. (1999). *Tracking Multiple Items Through Occlusion: Clues to Visual Objecthood*. *Cognitive Psychology* (Vol. 38). Retrieved from http://www.idealibrary.comon

Seed, A., & Byrne, R. (2010, December 7). Animal tool-use. *Current Biology*. Cell Press. https://doi.org/10.1016/j.cub.2010.09.042

Seed, A., & Tomasello, M. (2010). Primate cognition. *Topics in Cognitive Science*, *2*(3), 407–419. https://doi.org/10.1111/j.1756-8765.2010.01099.x

Shimizu, Y. A., & Johnson, S. C. (2004). Infants' attribution of a goal to a morphologically unfamiliar agent. *Developmental Science*, *7*(4), 425–430. https://doi.org/10.1111/j.1467-7687.2004.00362.x

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, *14*(1), 29–56. https://doi.org/10.1207/s15516709cog1401_3

Spelke, E. S. (1998). Nativism, empiricism, and the origins of knowledge. *Infant Behavior and Development*, *21*(2), 181–200. https://doi.org/10.1016/S0163-6383(98)90002-9

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

Spelke, E. S., & Kinzler, K. D. (2007, January). Core knowledge. *Developmental Science*. Blackwell Publishing Ltd. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Stavans, M., & Baillargeon, R. (2016). Four-month-old infants individuate and track simple tools following functional demonstrations. *Developmental Science*, (October 2015), 1–13. https://doi.org/10.1111/desc.12500

Stavans, M., Lin, Y., Wu, D., & Baillargeon, R. (2019). Catastrophic individuation failures in infancy: A new model and predictions. *Psychological Review*, *126*(2), 196–225. https://doi.org/10.1037/rev0000136

Surian, L., & Caldi, S. (2010). Infants' individuation of agents and inert objects.

Developmental Science, *13*(1), 143–150. https://doi.org/10.1111/j.1467-7687.2009.00873.x

Tao, R. (2016). *Understanding object-directed intentionality in Capuchin monkeys and humans*. University of St Andrews. https://doi.org/10.1002/ejoc.201200111

Taylor, A. H., Miller, R., & Gray, R. D. (2012). New Caledonian crows reason about hidden causal agents. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(40), 16389–16391. https://doi.org/10.1073/pnas.1208724109

Taylor, A. H., Miller, R., & Gray, R. D. (2013a). Reply to Boogert et al.: The devil is unlikely to be in association or distraction. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(4), E274. https://doi.org/10.1073/PNAS.1220564110

Taylor, A. H., Miller, R., & Gray, R. D. (2013b). Reply to Dymond et al.: Clear evidence of habituation counters counterbalancing. *Proceedings of the National Academy of Sciences*, *110*(5), E337–E337. https://doi.org/10.1073/pnas.1219586110

Tomasello, M., & Call, J. (1997). *Primate cognition.* Oxford University Press. Retrieved from http://psycnet.apa.org/psycinfo/1997-36461-000

Tomasello, M., & Herrmann, E. (2010). Ape and Human Cognition: What's the Difference? *Current Directions in Psychological Science*, *19*(1), 3–8. https://doi.org/10.1177/0963721409359300

Träuble, B., Pauen, S., & Poulin-Dubois, D. (2014). Speed and direction changes induce the perception of animacy in 7-month-old infants. *Frontiers in Psychology*, *5*(OCT), 1–8. https://doi.org/10.3389/fpsyg.2014.01141

Tsutsumi, S., Ushitani, T., Tomonaga, M., & Fujita, K. (2012). Infant monkeys' concept of animacy: The role of eyes and fluffiness. *Primates*, *53*(2), 113–119. https://doi.org/10.1007/s10329-011-0289-8

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

Vallortigara, G. (2012). Core knowledge of object, number, and geometry: A comparative and neural approach. *Cognitive Neuropsychology*, *29*(1–2), 213–236. https://doi.org/10.1080/02643294.2012.654772

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., Fragaszy, D. M., & Savage-Rumbaugh, S. (1995). Performance in a Tool-Using Task by Common Chimpanzees (Pan troglodytes), Bonobos (Pan paniscus), an Orangutan (Pongo pygmaeus), and Capuchin Monkeys (Cebus apella). *Journal of Comparative Psychology*, *109*(1), 52–60. https://doi.org/10.1037/0735-7036.109.1.52

Visalberghi, E., & Tomasello, M. (1998). Primate causal understanding in the physical and psychological domains. *Behavioural Processes*, *42*(2–3), 189–203. https://doi.org/10.1016/S0376-6357(97)00076-4

Völter, C. J., & Call, J. (2014). Great apes (Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo abelii) follow visual trails to locate hidden food. *Journal of Comparative Psychology*, *128*(2), 199–208. https://doi.org/10.1037/a0035434

Völter, C. J., & Call, J. (2017). Causal and inferential reasoning in animals. In *APA handbook of comparative psychology: Perception, learning, and cognition.* (pp. 643–671). American Psychological Association. https://doi.org/10.1037/0000012-029

Völter, C. J., Sentís, I., & Call, J. (2016). Great apes and children infer causal relations from patterns of variation and covariation. *Cognition*, *155*, 30–43. https://doi.org/10.1016/j.cognition.2016.06.009

Vonk, J., & Povinelli, D. J. (2012). Similarity and Difference in the Conceptual Systems of PrimatesThe Unobservability Hypothesis. In *The Oxford Handbook of Comparative Cognition*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195392661.013.0029

Vonk, J., & Subiaul, F. (2009). Do chimpanzees know what others can and cannot do? Reasoning about “capability.” *Animal Cognition*, *12*(2), 267–286. https://doi.org/10.1007/s10071-008-0189-5

Wang, S. hua, & Baillargeon, R. (2008). Detecting impossible changes in infancy: a three-system account. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2007.10.012

Weiss, D. J., & Santos, L. R. (2006). *Why Primates? The Importance of Nonhuman Primates for Understanding Human Infancy*. *INFANCY* (Vol. 9).

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., Alexander, G. M., Wheeler, L., & Norvell, J. M. (2012). Sex differences during visual scanning of occlusion events in infants. *Developmental Psychology*, *48*(4), 1091–1105. https://doi.org/10.1037/a0026529

Wilcox, T., & Baillargeon, R. (1998a). 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., & Baillargeon, R. (1998b). Object individuation in young infants: Further evidence with an event-monitoring paradigm. *Developmental Science*, *1*(1), 127–142. https://doi.org/10.1111/1467-7687.00019

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., & Chapa, C. (2004). Priming infants to attend to color and pattern information in an individuation task. *Cognition*, *90*(3), 265–302. https://doi.org/10.1016/S0010-0277(03)00147-1

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

Wilcox, T., Woods, R., & Chapa, C. (2008). Color–function categories that prime infants to use color information in an object individuation task. *Cognitive Psychology*, *57*, 220–261. https://doi.org/10.1016/j.cogpsych.2008.02.001

Wood, J. N., Glynn, D. D., & Hauser, M. D. (2008). Rhesus monkeys' understanding of actions and goals. *Social Neuroscience*, *3*(1), 60–68. https://doi.org/10.1080/17470910701563442

Wood, J. N., & Hauser, M. D. (2008). Action comprehension in non-human primates: motor simulation or inferential reasoning? *Trends in Cognitive Sciences*, *12*(12), 461–465. https://doi.org/10.1016/J.TICS.2008.08.001

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4

Wu, Y., Muentener, P., & Schulz, L. E. (2016). The Invisible Hand: Toddlers Connect Probabilistic Events With Agentive Causes. *Cognitive Science*, *40*(8), 1854–1876. https://doi.org/10.1111/cogs.12309

Xu, F. (1999). Object individuation and object identity in infancy: the role of spatiotemporal information, object property information, and language. *Acta Psychologica*, *102*(2–3), 113–136. https://doi.org/10.1016/S0001-6918(99)00029-3

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., & Baker, A. (2005). Object Individuation in 10-Month-Old Infants Using a Simplified Manual Search Method. *Journal of Cognition and Development*, *6*(3), 307–323. https://doi.org/10.1207/s15327647jcd0603_1

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

Xu, F., Carey, S., & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive Psychology*, *49*(2), 155–190. https://doi.org/10.1016/j.cogpsych.2004.01.001

Xu, F., Carey, S., & Welch, J. (1999). Infants' ability to use object kind information for object individuation. *Cognition*, *70*(2), 137–166. https://doi.org/10.1016/S0010-0277(99)00007-4

Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science*, *16*(5), 372–377. https://doi.org/10.1111/j.0956-7976.2005.01543.x