



## **Are monkeys able to discriminate appearance from reality?**

Marie Hirel, Constance Thiriau, Inès Roho, Hélène Meunier

### **► To cite this version:**

Marie Hirel, Constance Thiriau, Inès Roho, Hélène Meunier. Are monkeys able to discriminate appearance from reality?. *Cognition*, 2020, 196, pp.104123 -. [10.1016/j.cognition.2019.104123](https://doi.org/10.1016/j.cognition.2019.104123). [hal-03488408](https://hal.science/hal-03488408)

**HAL Id: hal-03488408**

**<https://hal.science/hal-03488408v1>**

Submitted on 21 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

## Are monkeys able to discriminate appearance from reality?

HIREL Marie<sup>a,b</sup> – THIRIAU Constance<sup>a,b</sup> – ROHO Inès<sup>a,b</sup> – MEUNIER Hélène<sup>a,b</sup>

<sup>a</sup> Centre de primatologie de l'Université de Strasbourg, Fort Foch 67207, Niederhausbergen, France

<sup>b</sup> Laboratoire de Neurosciences Cognitives et Adaptatives, UMR 7364, Université de Strasbourg, France

Corresponding author

- E-mail address: hirel.marie@gmail.com
- Postal address: Centre de Primatologie de l'Université de Strasbourg, Fort Foch 67207, Niederhausbergen, France

### Abstract:

The understanding that the perceptual appearance of the environment can differ in several ways from the reality underlies the ability to discriminate appearance from reality. Being able to realize when a misperception can lead us to behave in inappropriate ways confers an evolutionary advantage and may be a prerequisite to develop a Theory of Mind. Understanding that our own perception can differ from reality seems indeed necessary to attribute to others perceptions or beliefs different than ours. This appearance-reality discrimination ability has recently been demonstrated in great apes but no information is currently available regarding this ability in other nonhuman species. In a comparative study, we tested Tonkean macaques (*Macaca tonkeana*), an Old World primate species, and brown capuchins (*Sapajus apella*), a New World primate species. We provided monkeys with two experiments using visual illusions of size and quantity to test their ability to discriminate appearance from reality, with an experimental setup similar to the one developed by Krachun and colleagues (2016) on chimpanzees. A large number of brown capuchins, from different ages and both sexes, as well as two Tonkean macaques succeeded in the two experiments. By ruling out all alternative explanations (i.e. visual tracking or associative learning), our study brings the first evidence that some Old World and New World monkeys are able to discriminate appearance from reality. Our results suggest moving the evolutionary apparition of this cognitive ability earlier in time. Finally, it suggests that humans could share more Theory of Mind components with more nonhuman species than we previously thought.

Key words: appearance-reality discrimination, monkeys, brown capuchins (*Sapajus apella*), Tonkean macaques (*Macaca tonkeana*), theory of mind.

## 1. Introduction

Perception is an essential first step for animals to gather information about their physical and social environment. It is necessary for individuals to properly perceive their surroundings in order to adjust their behaviour and to anticipate environmental changes. Most of the time, perception is reliable in terms of providing information about the way the world is. However, animals are confronted every day with situations in which things appear different from what they really are (Matsuno & Fujita, 2009; Moll & Tomasello, 2012). They can be misled by information they receive from objects, individuals, actions, or even experiences. These situations, in which things are not what they appear to be, can result from simple perceptual illusions, as for example when a straight stick in water looks bent, or mistaking a snake for a tree branch. But being fooled can also happen in more complex social interactions, such as among human beings when a person intentionally misleads another through the use of lies or other disguises (Flavell, Flavell, & Green, 1983; Krachun, Lurz, Russell, & Hopkins, 2016). The ability to avoid mistakes in judgement based on misperception is known as appearance-reality (AR) discrimination. This distinction between appearance and reality has a large ecological importance. According to Flavell and colleagues (1986, page 1), it “assumes many forms, arises in many situations and can have serious consequences for our lives. The relation between appearance and reality figures importantly in everyday perceptual, conceptual, emotional and social activity – in misperceptions, misexpectations, misunderstandings, false beliefs, deception, play, fantasy, and so forth”. Thereby, being able to realize when this misperception can lead us to misconstrue our environment and thus to behave in inappropriate ways confers an obvious evolutionary advantage (Flavell et al., 1983; Karg, Schmelz, Call, & Tomasello, 2014; Krachun et al., 2016; Moll & Tomasello, 2012).

The ability to discriminate AR in social context could be considered as a crucial component of the human cognitive skill set known as “Theory of Mind” (ToM), i.e. ability to attribute mental states to oneself and to others, like intentions, beliefs, knowledge or perceptions (Premarck & Woodruff, 1978). ToM is often described as constituted by different cognitive components such as the ability to follow the gaze of others, attention-reading, intention-reading, perspective-taking, or the ability to attribute false-beliefs to others (Call & Tomasello, 2008; Meunier, 2017; Premarck & Woodruff, 1978). Having this ability appears to be a huge advantage in competition, in cooperation and even for all forms of communication with conspecifics (Baron-Cohen, 1989). According to some researchers, understanding others’ minds may require using one’s own experiences. Attributing mental states to others seems to be more complex than dealing with our own mental states, so self-understanding is more likely to develop first (Gopnik & Astington, 1988; Krachun, Call, & Tomasello, 2009; Wellman et al., 2001). Thus, understanding that our own perception could differ from reality seems necessary to attribute different perceptions than ours. These arguments

36 coincide with the idea that AR discrimination may be a prerequisite for the development of a fully  
37 formed ToM.

38 An AR research program was first introduced to children by Braine & Shanks (1965), and  
39 was then developed deeper by Flavell and his colleagues in the 1980s (1983, 1986, 1987). Their AR  
40 tests merely consisted of presenting ambiguous objects to children and then questioning them about  
41 the appearance and reality of objects (“What does the object look like?”, “What is it really?”).  
42 Experimenters used identity illusions (e.g. rock-sponge test where a sponge is painted to look like a  
43 rock) or modified the properties of objects by using distorting lenses, colour filters or mirrors  
44 (Braine & Shanks, 1965; Flavell, 1993; Flavell et al., 1983; Flavell, Green, Flavell, Watson, &  
45 Campione, 1986; Flavell, Green, Wahl, & Flavell, 1987). Then, an array of studies has flourished  
46 with revised nonverbal tasks to remove unnecessary difficulties of verbal comprehension or word  
47 confusing. These revisited tests have still shown an increase of the AR ability with age. Although  
48 they make some mistakes, three-year-old children (and even some two and a half year old) mainly  
49 succeeded in those tests by distinguishing appearance from reality of objects (e.g. Hansen &  
50 Markman, 2005; Karg et al., 2014; Moll & Tomasello, 2012; Sapp, Kang, & Muir, 2000). The  
51 stepwise effect of age found in understanding AR supports the claims of a considerable cognitive  
52 improvement progressively throughout child development with some abilities developing before  
53 others (Wellman et al., 2001).

54 Meanwhile, some researchers have focused their interest on the evolutionary roots of this AR  
55 ability by testing nonhuman primates. Our closest relatives no doubt encounter perceptual  
56 ambiguities in their environments and are sensitive to perceptual illusions like humans (Kelley &  
57 Kelley, 2014; Matsuno & Fujita, 2009). However, only a few studies have emerged, mainly  
58 focusing on chimpanzees, and thus our knowledge about nonhuman primates’ understanding of this  
59 phenomenon is still sparse. Chimpanzees have demonstrated strong evidence that they act according  
60 to the actions and perceptions of others (Bräuer, Call, & Tomasello, 2007; Buttelmann, Carpenter,  
61 Call, & Tomasello, 2007; Call & Tomasello, 2008; Hare, Call, & Tomasello, 2001; Krupenye,  
62 Kano, Hirata, Call, & Tomasello, 2016; Meunier, 2017; Tomasello, Call, & Hare, 2003) and they  
63 seem to misrepresent themselves in order to deceive others (e.g. Hare, Call, & Tomasello, 2006;  
64 Hirata & Matsuzawa, 2001; Whiten & Byrne, 1988; Woodruff & Premack, 1979). Recent  
65 systematic studies have also shown that several monkey species tend to perform very much like  
66 great apes in some cognitive tasks and reveal impressive communicative and social abilities (Amici,  
67 Aureli, & Call, 2010; Martin & Santos, 2016; Meunier, 2017; Schmitt, Pankau, & Fischer, 2012).  
68 Great apes and monkeys have also demonstrated that they possess another component of ToM:  
69 visual perspective-taking (Bräuer, Call, & Tomasello, 2007; Bray, Krupenye, & Hare, 2014;  
70 Canteloup, Piroux, Poulin, & Meunier, 2016; Flombaum & Santos, 2005; Hare, Addessi, Call,

71 Tomasello, & Visalberghi, 2003; Meunier, 2017; Overduin-de Vries, Spruijt, & Sterck, 2014). This  
72 ability may be positively correlated with performances found in children's AR tests (Flavell et al.,  
73 1986). Thus, investigating AR understanding in nonhuman primates is important to inform us about  
74 their cognitive abilities and on the evolution of human-like ToM. So far, three studies have tested  
75 AR discrimination in nonhuman animals but only in great apes, by adapting classic AR tests used  
76 with children, i.e. presenting subjects with a situation in which an objects' real and apparent  
77 properties were in conflict. Krachun and her colleagues (2009) carried out the first investigation by  
78 testing chimpanzees with a grape size illusion using distorting lenses: out of the 14 subjects tested,  
79 four of them were successful in the test. The authors have then decided to refine and replicate their  
80 study, in order to investigate the scope and flexibility of chimpanzees' AR discrimination, this time  
81 testing different kinds of visual illusions: grape size illusion using distorting lenses, quantity of  
82 grapes illusion using a mirror and colour illusion using coloured filters (Krachun et al., 2016). Their  
83 results showed that in the three illusion tests, some chimpanzees were able to adapt their behaviour  
84 by choosing the truly largest food reward, providing the first evidence that chimpanzees seem to  
85 possess the ability to discriminate appearance from reality. Furthermore, Karg and colleagues  
86 (2014) used the same paradigm to test children and great apes on a size illusion too, by occluding  
87 portions of one large and one small food stick. They reached the same conclusion: like children,  
88 gorillas (*Gorilla gorilla*), bonobos (*Pan paniscus*), orang-utans (*Pongo abelii*), and chimpanzees  
89 (*Pan troglodytes*) succeeded at group level. This comparative study then highlights that all great  
90 apes seem able to discriminate appearance from reality, however, our understanding of AR  
91 discrimination in more phylogenetically distant species is still lacking.

92 Our study was to investigate whether other nonhuman primates' species, more  
93 phylogenetically distant to humans than great apes, have the ability to distinguish appearance from  
94 reality. To investigate this, we conducted a comparative study testing both an Old World monkey  
95 species, Tonkean macaque (*Macaca tonkeana*), and a New World monkey species, brown capuchin  
96 (*Sapajus apella*). Choosing these species should allow us to go further in evolution than the  
97 common ancestor of great apes (about 15 million years ago) and provide better insights into the  
98 evolutionary roots of AR understanding in primates. If both species appear able to discriminate AR,  
99 it would imply that this ability emerged earlier than, or with, the common ancestor of *Sapajus* and  
100 *Homo* (over 35 million years ago). In contrast, if macaques but not capuchins succeed in these AR  
101 tests, it would suggest that this ability emerged with the common ancestor of *Macaca* and *Homo*  
102 (over 25 millions years ago). Finally, if capuchins succeed in these experiments but macaques fail,  
103 convergent evolution of this ability could be considered. To be able to reliably compare monkeys'  
104 performances with those of great apes, we provided monkeys with a procedure very similar to the  
105 one developed by Krachun and colleagues (2016) in their study on chimpanzees. Moreover, to rule

out all low-level alternative explanations of potential success, we added an additional control procedure. Several macaques and capuchins species, particularly Tonkean macaques and brown capuchins, have already demonstrated some complex cognitive abilities of ToM, like gaze following (e.g. Amici, Visalberghi, & Call, 2009; Tomasello, Call, & Hare, 1998), attentional reading (e.g. Canteloup, Bovet, & Meunier, 2015; Defolie, Malassis, Serre, & Meunier, 2015), or perspective-taking ability (e.g. Bray, Krupenye, & Hare, 2014; Flombaum & Santos, 2005; Overduin-de Vries, Spruijt, & Sterck, 2014; Canteloup et al., 2016; Hare, Addessi, Call, Tomasello, & Visalberghi, 2003). Thereby, we could expect that these two species might succeed in our AR understanding experiments.

115

## 116 **2. Materials and methods**

### 117 2.1. Subjects and experimental conditions

118 This study was conducted at the Centre de Primatologie de l'Université de Strasbourg,  
119 France. Data was collected from March to June 2018. All subjects of both species were living in  
120 stable social groups. The group of Tonkean macaques (*Macaca tonkeana*) were composed of 28  
121 individuals (12 females and 16 males) from one to 22 years old, living in a wooded outdoor  
122 enclosure of 3788 m<sup>2</sup> with constant access to an indoor room of 20 m<sup>2</sup>. The group of brown  
123 capuchins (*Sapajus apella*) were composed of 20 individuals (11 females, 8 males and one non-  
124 sexed baby) from one month to 30 years old, living in a wooded outdoor enclosure of 2332 m<sup>2</sup> with  
125 constant access to an indoor room of 16.21 m<sup>2</sup>. All animals were fed and had access to water *ad*  
126 *libitum*. Each subject was tested individually by sitting in front of the experimenter at a table,  
127 situation in an experimental room next to their outdoor enclosure. Experimental devices were  
128 displayed on the table and individuals were required to make a choice by pointing to one of them.  
129 These two groups had previously participated in several ethological studies, but they had no  
130 previous experience with AR discrimination tasks. All animals participated in experiments by  
131 entering the experimental room voluntarily and thus accepting to be isolated from the rest of the  
132 group. Thereby, eight Tonkean macaques (seven males and one female; see Table 1) and eleven  
133 brown capuchins (seven males and four females; see Table 1) participated in our experiments. In  
134 addition, this study respected the European ethical standards and regulations (directive  
135 2010/63/EU).

136

### 137 2.2. Lens experiment

138 The aim of the lens experiment was to investigate the ability of subjects to recognize when the  
139 size of an object seemed different than it actually is. To produce a size illusion, we used exactly the  
140 same devices as those used in Krachun and colleagues (2016): two identical boxes with distorting

lenses (magnifying and minimizing) mounted into their front side. A small grape was placed in the magnifying box just behind the lens to appear larger, and a large grape was placed in the minimizing box just behind the lens, so it appeared smaller (see Figure 1). The sizes of the small and large grapes (2x1.5 cm and 2.75x1.75 cm, respectively) were the same in all tests. This experiment included five successive tests: the preference test, the basic test, the seen tracking test, the unseen tracking test and the avoid-lens test. Each one of them (except the preference test) was composed of several sessions containing two or four warmup preference trials followed immediately by two demo trials and then 12 experimental trials. The individual needed to succeed in one test in order to move on to the next one, meaning that he had to choose the truly larger grape significantly more often than 50% chance within the maximum number of sessions allowed. If not, they were removed from the subsequent experiments. In each test, the position of the large grape was randomised across trials: it was on the left and right sides an equal number of times per session and was never in the same position for more than two consecutive trials. We ran a maximum of one session per individual per half-day, and at least two sessions in each test (except for the preference test), in order to ensure a familiarisation of subjects with tasks. Unlike the Tonkean macaques, the brown capuchins were less familiar with the experimental room and could not be isolated from their group during one whole session. Thus, we decided to test them with half-sessions at a time (two-four warmup preference trials, two demo trials and six trials), with a maximum of two half-sessions per half-day.

### *2.2.1. Preference test*

With the preference test, we wanted to test the preference of each subject for larger over smaller grapes. It was a simple choice test, consisting of a maximum of two sessions of 12 trials each, between a small and a large grape displayed on the table by the experimenter in front of the animal. If the subject succeeded within the maximum of the two sessions allowed, we proceeded to the basic test. Moreover, as a warmup to the testing sessions and to confirm the individual's motivation, we ran two preference trials before each following session. If the individual chose the large grape in both of these trials, we moved on with the session. If the individual chose the large grape only one time among the two trials, we added two more warmup preference trials and then the individual had to choose the large grape at least 3/4 trials to proceed to the session. If he did not, the session was delayed.

### *2.2.2. Demo trials*

Since these animals had no previous experience with distorting lenses, two demo trials were carried out to allow them to learn about their effect, just after the warmup preference trials in each

176 following session. As the individual was watching, the magnifying and minimizing boxes were  
177 placed side by side in the middle of the table. The small grape was put behind the magnifying lens  
178 and the large grape behind the minimizing lens. This way, the large grape appeared smaller and the  
179 small grape appeared larger. The experimenter then moved boxes forward and apart on the table, so  
180 the subject could make its choice. Once the choice was made, she (the experimenter) removed the  
181 lenses from the boxes several times to demonstrate the transformation of the grapes' appearance and  
182 show the proprieties of lenses. After this demonstration, she gave the chosen grape to the individual  
183 and discarded the other.

### 184 185 2.2.3. *Basic test*

186 The goal of the basic test was to examine if individuals were able to ignore the deceptive size  
187 appearance of grapes and so if they were capable to choose the truly bigger one. The procedure was  
188 similar to the demo trials except that we did not demonstrate the lenses' effect, i.e. after the  
189 individual had made his choice, the experimenter removed the lenses of the two boxes, gave the  
190 chosen grape and discarded the other one. If individuals succeeded in this test within a maximum of  
191 eight sessions, they could proceed to the tracking control test.

### 192 193 2.2.4. *Tracking control tests*

194 The tracking control test was carried out in order to rule out the possibility that individuals  
195 succeeded in the basic test by visually tracking the truly large grape. Thus, to ensure that they were  
196 making an AR discrimination, two tracking control tests were added.

#### 197 198 2.2.4.1. Seen tracking test

199 The purpose of the seen tracking control test was to familiarize individuals with the new  
200 procedure of stacking boxes and was thus a transition to the unseen tracking test. The experimenter  
201 placed the two boxes on the table and baited them with their corresponding grapes, then she stacked  
202 the boxes in the middle of the table. While the subject was still watching, she distributed the boxes  
203 on each side and moved them forward and apart to let the individual choose. The side and top-  
204 bottom positions of boxes were randomised across trials. Individuals had up to a maximum of four  
205 sessions to succeed in this test to move on to testing.

#### 206 207 2.2.4.2. Unseen tracking test

208 The procedure of the unseen tracking test was similar to the seen tracking test: the stacking  
209 and the baiting of the boxes were done in sight of the individuals. But before unstacking the two  
210 boxes and distributing them on left and right sides for the subjects to make their choice, an opaque



occluder was placed so that the subject could not see where the large grape was placed. Then the experimenter removed the occluder and moved the boxes forward and apart. In this way, it was impossible to succeed in this test by visually tracking the large grape. This procedure compels individuals to remember that the smaller grape that appeared through the lens is truly the large one, and *vice versa*. Individuals who passed this test within a maximum of four sessions proceeded to the avoid-lens control test.

### 2.2.5. *Avoid-lens control tests*

The avoid-lens control test was a final control put in place to rule out the possibility that our subjects succeeded in all previous tests by learning to identify and avoid the magnifying lens, or by learning to identify and look for the minimizing lens, or even by merely learning to look for an apparent specific size of grape and choose it. In Krachun and colleagues' (2016) study on chimpanzees, the avoid-lens control test ruled out several potential learnings but their subjects could still succeed by avoiding the apparent large-sized grape. Thereby, the avoid-lens test n°1 is an adapted version of their learning control, in order to rule out a supplemental learning alternative explanation of a success previously mentioned. After this first learning control test, we decided to carry out the avoid-lens test n°2 which replicates thoroughly the avoid-lens test designed by Krachun and colleagues (2016), in order to truly compare performances of our subjects with those of chimpanzees. The avoid-lens test n°2 was carried out even on our subjects that failed the avoid-lens test n°1.

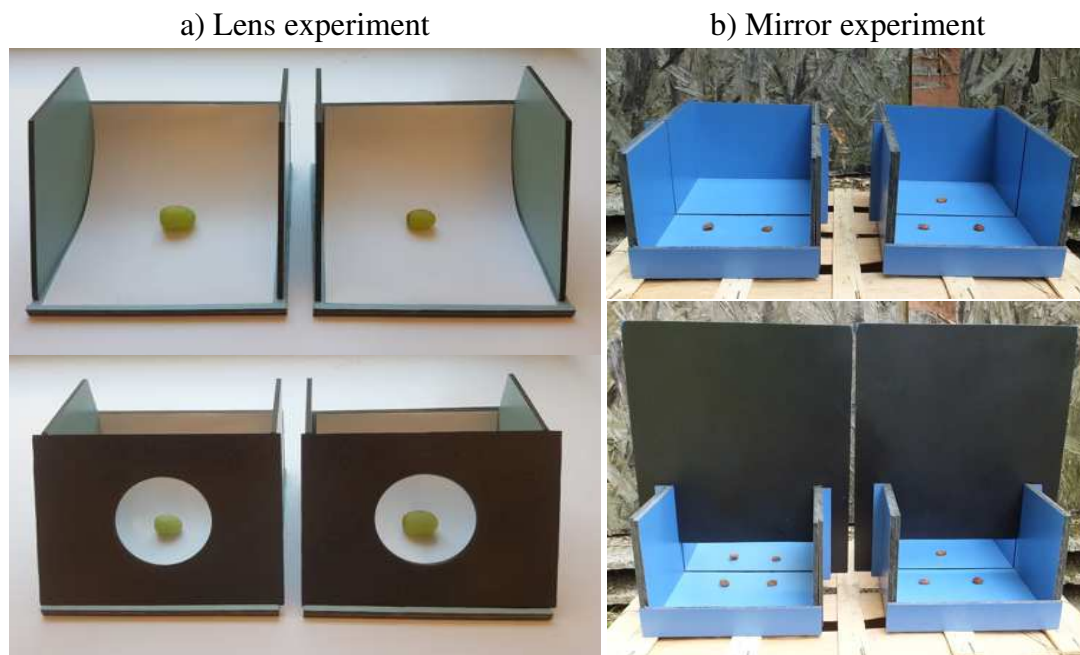
#### 2.2.5.1. Avoid-lens test n°1

In the avoid-lens test n°1, the experimenter used a magnifying lens and a regular non-distorting lens. The procedure was the same as the unseen tracking test except that the large grape was placed behind the magnifying lens to make it appear even larger, and the small grape was placed behind the regular lens (its apparent size corresponding to its real size). In this situation, subjects had to choose, and not avoid, the magnifying box to obtain the truly larger grape. They also had to give up on the grape which looked the same size as that which had been correct in previous trials (i.e. size of the small grape) and choose the other one instead. We gave our subjects a maximum of four sessions to succeed in this test.

#### 2.2.5.2. Avoid-lens test n°2

The procedure of the avoid-lens test n°2 was the same as the avoid-lens test n°1 except that the experimenter used a large grape and an even larger one. The large grape was placed behind the

regular lens and the even larger grape was placed behind the magnifying lens. Subjects had up to a maximum of four sessions to succeed in this last test.



*Figure 1 : Experimental devices used in the two experiments to create visual illusion of size (a) and of quantity (b).*

### 2.3. Mirror experiment

The mirror experiment aimed at testing AR discrimination of our subjects with a quantity illusion. To test this, we used two identical rectangular boxes into which a mirror could be inserted in the middle (see Figure 1; all devices' dimensions were exactly the same as in the study of Krachun and colleagues (2016)). Groups of two or three raisins were placed in each box, in front of the mirror. In this way, the raisins in the mirror box appeared to double in number. To make the two boxes completely identical, vertical "collars" were slid down over the middle of each box in order to obscure the individual's view of the back section of the boxes and the mirror's top edge (see Figure 1). This experiment proceeded similarly as the Lens experiment with the same five successive tests: the preference test, the basic test, the seen tracking test, the unseen tracking test and the avoid-mirror test. Each one of them (except the preference test) was composed of several sessions containing two or four warmup preference trials, two demo trials and 12 trials. The individual needed to succeed in one test in order to move on to the next one, meaning that he had to choose the truly larger number of raisins significantly more often than 50% chance within the maximum number of sessions allowed. If not, they were removed from the subsequent experiments. The position of the larger number of raisins was randomised across trials: it was on the left and right side an equal number of times per session and was never in the same position for more than two consecutive trials. We ran a maximum of one session per subject per half-day for Tonkean

macaques. As for the Lens experiment, we tested brown capuchins with half-sessions at a time (two-four warmup preference trials, two demo trials and six trials), with a maximum of two half-sessions per half-day. For each species, we ran at least two sessions in each test (except for the preference test), in order to ensure a familiarisation of subjects with the task.

### *2.3.1. Preference test*

The preference test was designed to test the preference of each individual for larger over smaller numbers of raisins. It was a simple choice test in which a small and a large number of raisins were displayed in the boxes on the table. In the following experimental tests, individuals encountered the choice between two types of number combinations: 2 vs 3 raisins that appeared like 3 vs 4 raisins due to the mirror. Thus, we carried out two preference tests, each one composed of a maximum of four sessions of 12 trials: the first test was a choice between 2 vs 3 raisins, and the second one was a choice between 3 vs 4 raisins. If the subject succeeded in each preference test within their maximum of four sessions allowed, we proceeded to the basic test. Furthermore, as a warmup to testing sessions and to confirm the individual's motivation, we ran two preference trials before each following testing sessions (the number combination corresponded to how raisins will appear, i.e. 3 vs 4). If the individual chose the larger number of raisins in both of these trials, we moved on with the session. If the individual chose the larger number of raisins only one time among the two trials, we added two more warmup preference trials and then the individual had to choose the larger number of raisins at least 3/4 trials to proceed to the session. If not, the session was delayed.

### *2.3.2. Demo trials*

Since our subjects had no previous experience with mirrors, these demo trials were carried out to give them the opportunity to learn about their effect, just after the warmup preference trials in each following session. While the individual was watching, the two boxes were placed side by side on the table. The two groups of raisins (2 vs 3) were placed in the boxes and the mirror was inserted into the one containing fewer raisins. In this way, the two raisins appeared as a group of four. The experimenter then slid the mirror up and down several times to show to the subject how it made raisins appear to double in number. She then positioned the "collars" on the boxes and moved the boxes forward and apart on the table. Once the individual had made their choice, she executed a few actions to show the illusory properties of the mirror: removing the "collars" in order to make the mirror visible again, slowly sliding the mirror up and down at least twice, sweeping the hand through the non-mirror box to show that it did not contain a mirror, tilting the mirror box forward so that the individual could see that it was empty at the back, and finally removing the mirror and

315 placing it aside. After this demonstration, the chosen raisins were given to the subject and the others  
316 were discarded.

317

### 318 2.3.3. *Basic test*

319 The basic test aimed to examine if our subjects were able to optimize their reward by ignoring  
320 the illusion of a greater number of raisins and by choosing the truly larger group. The procedure  
321 was similar to the demo trials except that we did not demonstrate the effect of the mirror. As the  
322 subject was watching, the boxes were placed side by side on the table, the two groups of raisins (2  
323 vs 3) were positioned in each box, the mirror was inserted into the box containing fewer raisins,  
324 “collars” were slid down over the boxes, and finally the experimenter moved the boxes apart and  
325 forward. After the subject had made its choice, the experimenter removed the “collars”, followed by  
326 the mirror, and finally gave the chosen group of raisins to the subject, discarding the other one. If  
327 the subjects succeeded within a maximum of eight sessions, they could proceed to the tracking  
328 control test.

329

### 330 2.3.4. *Tracking control tests*

331 As in the Lens experiment, the tracking control test was carried out to rule out the possibility  
332 that these animals succeeded in the basic test by visually tracking the larger number of raisins.  
333 Thus, we ran two tracking control tests: the seen tracking test to get them used to the new stacking  
334 procedure and the unseen tracking test.

335

#### 336 2.3.4.1. *Seen tracking test*

337 For the seen tracking test, the experimenter first placed the two groups of raisins on the table.  
338 Then, she positioned one box in the middle of the table and inserted the larger number of raisins  
339 (i.e. 3) into it, stacked the other box and baited it with the smaller number of raisins (i.e. 2). While  
340 the individual was still watching, she slid the mirror into the top box followed by the “collar” and  
341 unstacked the boxes to slide the “collar” onto the other box. Finally, she moved the boxes apart and  
342 forward. The side of the box into which the raisins were put in as well as the positioning of the  
343 boxes on the table were randomised across trials. Once the subject had chosen, the “collars” were  
344 removed, followed by the mirror, after which the experimenter gave the chosen group of raisins to  
345 the subject and threw away the other one. Individuals had up to a maximum of four sessions to  
346 succeed in this test.

347

#### 2.3.4.2. Unseen tracking test

The procedure of the unseen tracking test was similar to the seen tracking test: the stacking and the baiting of the boxes were conducted in sight of the subjects. However, after inserting the mirror, the experimenter inserted the opaque occluder. In this way, the individual could not see the unstacking of the boxes, the placement of the “collars” and the positioning of the boxes on each side of the table. The opaque occluder was then removed and the boxes were pushed forward. Once the individual had chosen, the actions were carried out exactly the same as the seen tracking test. Subjects who passed this test within a maximum of four sessions proceeded to the avoid-mirror control test.

#### 2.3.5. Avoid-mirror test

The avoid-mirror test was a final control to rule out the possibility that the animals succeeded in all previous tests by learning to identify and avoid the mirror, or by learning to look for it and to avoid a specific number configuration of raisins. The experimenter now used a group combination of 1 vs 2 raisins. The procedure was the same as the unseen tracking test except that the single raisin was placed into the non-mirror box and the two raisins were placed into the mirror box so that it looked like a choice between 1 vs 4 raisins. In this situation, individuals had to choose (and not avoid) the mirror box, and what appeared like four raisins, to obtain the truly larger quantity, of which had been the incorrect choice throughout all the previous trials. We gave to the subjects a maximum of four sessions to succeed in this test.

### 3. Results

Due to the small number of subjects that progressed to the final tests (and the exclusion of those which didn't), we first focused on individual performances. In both experiments, the binomial test was used to test if subjects significantly chose one type of stimulus more than the other (the larger or the smaller grape in the lens test, the larger or the smaller number of raisins in the mirror test). To pass a test, each subject had to choose the truly larger grape or the truly larger number of raisins in at least 10/12 trials in one session (exact binomial test:  $N=12$ ,  $p \leq 0.019$ ) or at least 8/12 trials in each of two consecutive sessions (exact binomial test:  $N=24$ ,  $p \leq 0.038$ ). As we ran at least two sessions in each test, a subject that obtained a score of 10/12 in the first session still had to choose the larger food reward at least 8/12 in the second session. If not, the subject had to complete another session with significant scores to move on to the next test. For subjects showing no preference for one type of stimulus before the control tests, i.e. in basic test and seen tracking test, a binomial test was run to compare their number of choices on each side of the setup, to rule out the possibility of a pointing bias as an explanation of individuals' failure. Our procedure of progressive

session generates an increasing probability of false positives. The cumulative type I error probability for reaching the criterion is already 0.053 in the first session and 0.097 in the second one. Hence, we tested group performances by combining individual performances corrected for multiple testing into a collective one, using Fisher's omnibus test (Quinn & Keough, 2002). To control for potential experimenter bias, all trials were videotaped and 20% of them was scored by a second blind coder (contingency coefficient:  $k=0.97$ ).

Species	Subjects	Sex	Age (years, months)	Preference test		Basic test								Seen tracking test				Unseen tracking test				Avoid-lens test n°1				Avoid-lens test n°2			
				S1	S2	S1	S2	S3	S4	S5	S6	S7	S8	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4
Tonkean macaques	Abrirot	M	4, 10	9	<u>11</u>	7	<u>8</u>	<u>8</u>	.	.	.	.	.	7	<u>8</u>	<u>9</u>	.	<u>8</u>	<u>9</u>	.	.	7	5	8	<u>11</u>	9	<u>10</u>	.	.
	Anubis	M	4, 9	<u>11</u>	.	10	<u>11</u>	.	.	.	.	.	.	12	<u>10</u>	.	.	10	<u>12</u>	.	.	4	6	4	8	12	<u>12</u>	.	.
	Yannick	M	8, 4	<u>11</u>	.	12	<u>12</u>	.	.	.	.	.	.	11	<u>11</u>	.	.	<u>9</u>	<u>9</u>	.	.	6	5	6	6	<u>8</u>	<u>8</u>	.	.
	Nereis	F	19, 2	<u>12</u>	.	<u>9</u>	<u>9</u>	.	.	.	.	.	.	<u>10</u>	<u>9</u>	.	.	<u>9</u>	<u>9</u>	.	.	7	6	8	6				
	Olli	M	7, 3	9	<u>10</u>	9	6	<u>10</u>	.	.	.	.	.	8	7	<u>10</u>	.	9	<u>12</u>	.	.	6	5	3	5	7	4	4	5
	Walt	M	10, 4	8	<u>11</u>	9	<u>11</u>	.	.	.	.	.	.	8	6	9	<u>10</u>	6	8	<u>10</u>	.	5	8	6	7				
	Olaf	M	6, 1	8	<u>11</u>	6	6	7	6	8	6	9	<u>10</u>	11	<u>11</u>	.	.	9	<u>12</u>	.	.	1	3	2	1	8	6	5	5
	Alvin	M	5, 1	9	<u>10</u>	4	6	4	7	9	7	<u>10</u>	.	6	7	8	<u>10</u>	12	<u>12</u>	.	.	2	0	5					
Brown capuchins	Kiri	M	5, 11	<u>10</u>	.	<u>9</u>	<u>9</u>	.	.	.	.	.	.	6	<u>10</u>	.	.	8	<u>10</u>	.	.	<u>8</u>	<u>8</u>	.	.	10	<u>12</u>	.	.
	Brugnon	M	3, 4	<u>9</u>	<u>9</u>	11	<u>12</u>	.	.	.	.	.	.	9	<u>11</u>	.	.	8	<u>10</u>	.	.	12	<u>12</u>	.	.	8	<u>10</u>	.	.
	Willow	F	10, 7	<u>11</u>	.	11	<u>10</u>	.	.	.	.	.	.	11	<u>12</u>	.	.	11	<u>11</u>	.	.	7	<u>11</u>	.	.	10	<u>10</u>	.	.
	Franklin	M	7, 7	<u>12</u>	.	12	<u>12</u>	.	.	.	.	.	.	<u>10</u>	<u>9</u>	.	.	<u>10</u>	<u>8</u>	.	.	5	8	<u>12</u>	.	10	<u>10</u>	.	.
	Balin	M	3, 9	<u>8</u>	<u>9</u>	12	<u>10</u>	.	.	.	.	.	.	10	<u>12</u>	.	.	12	<u>12</u>	.	.	6	8	6	<u>12</u>	10	<u>12</u>	.	.
	Popeye	M	16, 9	<u>12</u>	.	10	<u>10</u>	.	.	.	.	.	.	11	<u>12</u>	.	.	11	<u>12</u>	.	.	4	6	<u>9</u>	<u>9</u>	<u>9</u>	<u>9</u>	.	.
	Conan	M	2, 5	<u>12</u>	.	5	6	7	6	6	7	<u>8</u>	<u>8</u>	10	<u>11</u>	.	.	8	<u>12</u>	.	.	8	<u>11</u>	.	.	11	<u>12</u>	.	.
	Litchi	F	8, 10	7	<u>12</u>	9	<u>10</u>	.	.	.	.	.	.	11	<u>12</u>	.	.	12	<u>12</u>	.	.	4	4	7	8	5	6	6	6
	Bombers	M	3, 10	<u>9</u>	<u>8</u>	5	8	<u>11</u>	.	.	.	.	.	7	6	9	<u>10</u>	12	<u>12</u>	.	.	2	0	0	1	0	0	1	0
	Kinika	F	25, 9	<u>12</u>	.	8	7	<u>8</u>	<u>9</u>	.	.	.	.	6	8	5	6	.	.	.	.	.	.	.	.	.	.	.	.
	Capuche	F	2, 7	<u>10</u>	.	8	7	5	7	5	6	6	6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table 1: Number of correct trials (out of 12) in which individuals chose the truly larger grape in each session (S) for the Lens experiment. Underlined black scores correspond to the criterion of individuals' success of the test, i.e. choosing the large grape at least 10/12 in one session or 8/12 in two consecutive sessions.

### 3.1. Lens experiment

#### 3.1.1. Tonkean macaques

Eight subjects (seven males and one female; see Table 1) participated in the experiment. All subjects passed the preference test, three of which passed within the first session. All subjects passed the basic test in an average of 3.4 sessions, three of which passed in the first session. The eight subjects also succeeded in the seen tracking test within an average of 2.25 sessions, four of which passed in the first session. Then all of them passed the unseen tracking control test within an average of 1.9 sessions, two of which passed in the first session. One individual passed the avoid-lens test n°1, six individuals failed to meet passing criteria, and one individual did not complete the sessions due to lack of motivations. Concerning the avoid-lens test n°2, only five individuals were

able to tested. Among them, three succeeded within the two first sessions, including two individuals that had failed the avoid-lens test n°1 (see Table 1). Because of these conflicting performances between the two avoid-lens tests, we carried out two more sessions with these two individuals with the procedure of the unseen tracking and both of them failed, with a score of 0/12. At group level, performances of the Tonkean macaques were above chance in the basic test (Fisher's omnibus test:  $\chi^2=38.05$ ,  $df=16$ ,  $p=0.001$ ), the seen tracking test (Fisher's omnibus test:  $\chi^2=40.71$ ,  $df=16$ ,  $p=0.0006$ ) and the unseen tracking test (Fisher's omnibus test:  $\chi^2=45.76$ ,  $df=16$ ,  $p=0.0001$ ). In contrary, they were not above chance in the avoid-lens test n°1 (Fisher's omnibus test:  $\chi^2=3.95$ ,  $df=14$ ,  $p=0.996$ ) and avoid-lens test n°2 (Fisher's omnibus test:  $\chi^2=17.61$ ,  $df=10$ ,  $p=0.062$ ).

411

### 412 3.1.2. *Brown capuchins*

Eleven capuchins (seven males and four females; see Table 1) participated in the experiment. All subjects passed the preference test, seven of which passed within the first session. Ten subjects succeeded in the basic test in an average of 2.4 sessions, five of which passed in the first session. One female did not succeed in the basic test, and therefore was removed from subsequent tests. This female was found a significant left bias in pointing (exact binomial test:  $N=96$ ,  $p<0.0001$ ). Out of these 10 subjects, nine succeeded in the seen tracking test in an average of 1.5 sessions, six of which passed in the first session. One subject did not succeed in the seen tracking test and was found a significant left pointing bias (exact binomial test:  $N=96$ ,  $p<0.0001$ ). The nine remaining subjects also succeeded in the unseen tracking test with an average of 1.3 sessions, six of which passed in the first session. Then, seven capuchins passed the avoid-lens test n°1 in an average of 2.6 sessions, four of which passed within the two first sessions, and the two others failed. In the avoid-lens test n°2, five individuals succeeded in an average of 1.3 sessions, four of which passed in one session. At group level, performances of the brown capuchins were above chance in all tests (Fisher's omnibus test: basic:  $\chi^2=52.19$ ,  $df=22$ ,  $p=0.0003$ ; seen:  $\chi^2=50.91$ ,  $df=20$ ,  $p=0.0002$ ; unseen:  $\chi^2=52.83$ ,  $df=18$ ,  $p<0.0001$ ; avoid-lens n°1:  $\chi^2=36.04$ ,  $df=18$ ,  $p=0.007$ ; avoid-lens n°2:  $\chi^2=41.09$ ,  $df=18$ ,  $p=0.001$ ).

Species	Subjects	Sex	Age (years, months)	Preference test 2vs3				Preference test 3vs4				Basic test								Seen tracking test				Unseen tracking test				Avoid-mirror test			
				S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4	S5	S6	S7	S8	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4
Tonkean macaques	Abrirot	M	4, 10	<u>11</u>	.	.	.	<u>11</u>	.	.	.	10	<u>10</u>	.	.	.	.	.	.	3	8	<u>8</u>	.	9	5	<u>10</u>	.	12	<u>12</u>	.	.
	Olaf	M	6, 1	<u>12</u>	.	.	.	<u>12</u>	.	.	.	8	<u>11</u>	.	.	.	.	.	.	6	<u>10</u>	.	.	10	<u>9</u>	.	.	11	<u>12</u>	.	.
Brown capuchins	Kiri	M	5, 11	7	<u>10</u>	.	.	<u>10</u>	.	.	.	9	<u>9</u>	.	.	.	.	.	.	9	<u>10</u>	.	.	11	<u>9</u>	.	.	12	<u>12</u>	.	.
	Balin	M	3, 9	<u>10</u>	.	.	.	<u>10</u>	.	.	.	10	5	<u>10</u>	.	.	.	.	.	8	<u>10</u>	.	.	8	<u>8</u>	.	.	11	<u>11</u>	.	.
	Franklin	M	7, 7	<u>10</u>	.	.	.	<u>11</u>	.	.	.	10	<u>9</u>	.	.	.	.	.	.	10	<u>10</u>	.	.	9	<u>9</u>	.	.	12	<u>11</u>	.	.
	Popeye	M	16, 9	<u>11</u>	.	.	.	9	<u>11</u>	.	.	4	6	8	<u>9</u>	.	.	.	.	8	<u>10</u>	.	.	8	<u>9</u>	.	.	12	<u>9</u>	.	.
	Bombers	M	3, 10	9	<u>10</u>	.	.	<u>10</u>	.	.	.	8	<u>9</u>	.	.	.	.	.	.	8	<u>8</u>	.	.	8	<u>9</u>	.	.	12	<u>11</u>	.	.
	Litchi	F	8, 10	7	9	<u>9</u>	.	<u>10</u>	.	.	.	10	<u>11</u>	.	.	.	.	.	.	8	6	7	5	.	.	.	.	.	.	.	.

Table 2 : Number of correct trials (out of 12) in which individuals have chosen the truly larger number of raisins in each session (S). Underlined black scores correspond to the criterion of individuals' success of the test, i.e. choosing the larger number of raisins at least 10/12 in one sessions or 8/12 in two consecutive sessions.

### 3.2. Mirror experiment

#### 3.2.1. Tonkean macaques

Only two males Tonkean macaques participated in the experiment, because the others never came back to the experimental room or were no longer motivated to participate in the study (Table 2). Both individuals passed the two preference tests in a single session. Then, they succeeded in the basic test in two sessions, and in the seen and the unseen tracking test within maximum three sessions. Finally, they validated the avoid-mirror test within one session only. At group level, performances of these Tonkean macaques were above chance in all tests (Fisher's omnibus test: basic:  $\chi^2=11.74$ ,  $df=4$ ,  $p=0.019$ ; seen:  $\chi^2=10.54$ ,  $df=4$ ,  $p=0.032$ ; unseen:  $\chi^2=10.54$ ,  $df=4$ ,  $p=0.032$ ; avoid-mirror:  $\chi^2=11.74$ ,  $df=4$ ,  $p=0.019$ ).

#### 3.2.2. Brown capuchins

Six brown capuchins participated in the experiment (five males and one female; Table 2). These subjects passed the first preference test within a maximum of three sessions, and the second preference test within a maximum of two sessions. Then, all six succeeded in the basic test within an average of 2.2 sessions, four of which within two first sessions. Five succeeded in the seen tracking test in an average of 1.8 sessions, all of which passed within the first two sessions. One subject did not succeed the seen tracking test and thereby was removed from subsequent tests. This individual did not have a significant left or right bias in pointing (exact binomial test:  $N=48$ ,  $p=0.9$ ). The remaining five subjects succeeded in the unseen tracking test within an average of 1.8 sessions, all of which passed within the first two sessions. Finally, they validated the avoid-mirror test in only one session. At group level, performances of the brown capuchins were above chance in all tests



453 (Fisher's omnibus test: basic:  $\text{chisq}=32.09$ ,  $\text{df}=12$ ,  $p=0.001$ ; seen:  $\text{chisq}=29.35$ ,  $\text{df}=12$ ,  $p=0.003$ ;  
 454 unseen:  $\text{chisq}=29.35$ ,  $\text{df}=10$ ,  $p=0.001$ ; avoid-mirror:  $\text{chisq}=29.35$ ,  $\text{df}=10$ ,  $p=0.001$ ).  
 455

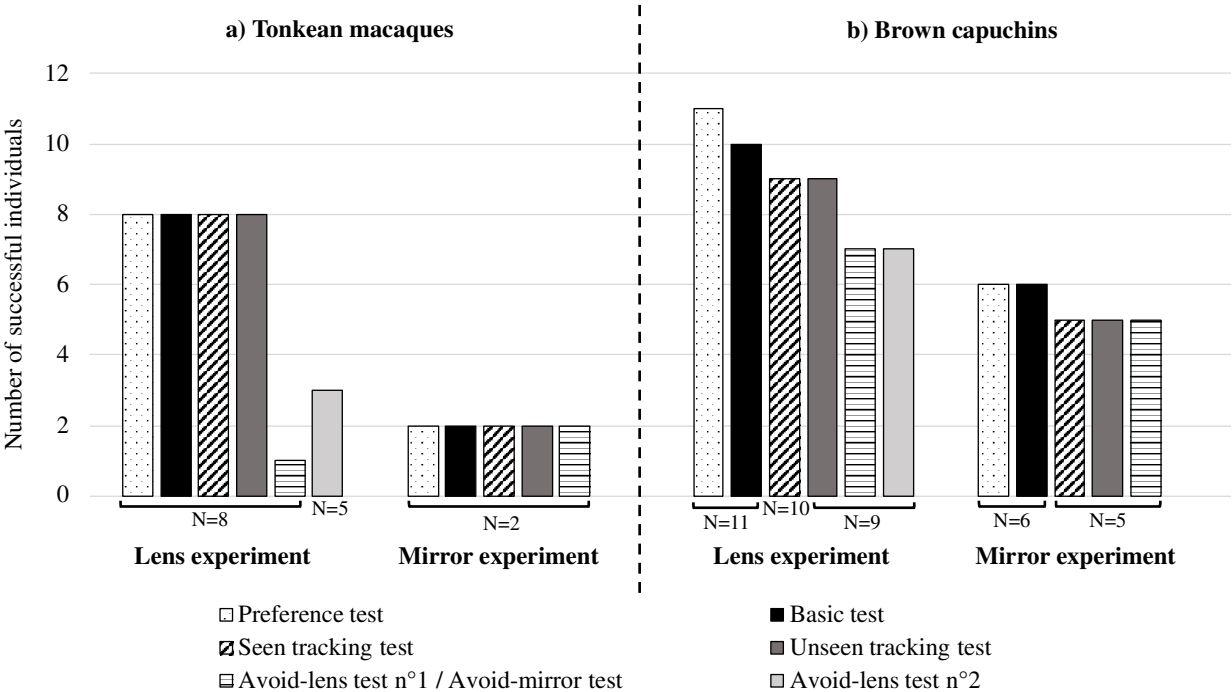


Figure 2 : Number of subjects of each species who passed each test of the Lens and Mirror experiment

456  
 457

#### 458 4. Discussion

459 This study brings the first evidence that monkeys are able to discriminate appearance from  
 460 reality. All our subjects passed the preference tests of both the Lens and Mirror experiments and  
 461 thus demonstrated a strong natural preference for larger or higher number of food items, as well as a  
 462 good size and quantity discrimination ability. Then, in the Lens experiment, all but two subjects of  
 463 each species passed successfully the basic test and the tracking control tests, demonstrating that  
 464 they were capable of ignoring the misleading appearance of the magnified grape to choose the truly  
 465 bigger one, without being able to visual track the grapes (see Figure 2). Afterwards, the large  
 466 number of successful brown capuchins, from different ages and both sexes, in the last learning  
 467 control brings strong evidence that they are able to discriminate appearance from reality. In  
 468 Tonkean macaques, our results are less clear, with only one individual succeeding in the avoid-lens  
 469 control (see Figure 2). Thus, we are unable to say with confidence whether or not this species  
 470 possesses appearance-reality discrimination ability. However, our procedure with lenses might have  
 471 not allow us to highlight AR discrimination ability in a larger scale in this species and could provide  
 472 results underestimating it (see methodological issues discussed below). This is supported by

473 Tonkean macaques' good performances in the Mirror experiment. Indeed, the two Tonkean  
474 macaques and all but one brown capuchin that were tested in this second experiment passed the  
475 visual tracking and the avoid-mirror tests. By validating all controls, they demonstrated they are  
476 able to discriminate perceived from real quantity.

477 We found that all except two individuals who were successful in the Mirror experiment had  
478 also succeeded in the Lens experiment. Successes in both of our experiments provides evidence for  
479 a general understanding in our animals, that perceptual appearance of objects can be different in  
480 several ways from reality. Nevertheless, we should approach this conclusion with caution for the  
481 few individuals that needed more than two sessions to succeed. Even by limiting the number of  
482 sessions allowed to reach a success, our and other procedures of progressive sessions generate an  
483 increasing probability of false positives. To account for this multiple testing issue, one option is to  
484 look out group performances. In this respect, brown capuchins succeeded significantly in all tests of  
485 the Lens and Mirror experiments. For Tonkean macaques, their succeeded significantly in all tests  
486 except for the avoid-lens tests n°1 and n°2. These results confirm our inability to conclude of an AR  
487 discrimination ability presence for Tonkean macaques with the Lens experiment. However, we still  
488 able to maintain confidence in our results demonstrating AR discrimination ability in both  
489 experiments for brown capuchins and in the Mirror experiment for Tonkean macaques. In summary,  
490 even if more subjects of both species are necessary to make a reliable conclusion at a species scale,  
491 results of the Mirror experiment confirmed those of the Lens experiment and together suggest that  
492 these monkeys could be able to discriminate appearance from reality.

493 The ability to discriminate appearance from reality was demonstrated in great apes by  
494 previous studies of Karg and colleagues (2014) and Krachun and colleagues (2009, 2016). Now, our  
495 study suggests the presence of this ability in Tonkean macaques, an Old World monkey species, and  
496 brings strong evidence in brown capuchins, a New World monkey species. Therefore, the  
497 hypothesis of the emergence of AR discrimination ability in a common ancestor dating back to at  
498 least the capuchin monkey genus is conceivable. However, our results certainly need to be  
499 strengthened by further research on more individuals and other species of macaques to make a more  
500 reliable conclusion. Since only two Tonkean macaques were compared to eight brown capuchins in  
501 both experiments, alternative hypotheses of convergent evolution cannot be completely ruled out.  
502 Thus, instead of looking at phylogenetic relatedness only, our interest should go towards the socio-  
503 ecological aspects that are characteristic of these species in order to improve our knowledge about  
504 the development of such cognitive abilities during evolution. Indeed, as example, many similarities  
505 were found between genus *Sapajus* and *Pan*: long life spans, explorative and manipulative  
506 behaviours, tool using, omnivorous diet, socially complex behaviours such as coalitions and  
507 cooperation and so forth (Visalberghi & McGrew, 1997).

508 One of the major aims of this study was to compare performances of monkey species with  
509 those of chimpanzees (Krachun et al., 2016) using a same experimental paradigm. Like  
510 chimpanzees, brown capuchins performed high scores in the two AR experiments and passed each  
511 test within a small amount of sessions. Not all of them succeeded but the proportion of successful  
512 brown capuchins is still greater than observed in chimpanzees. Hence, the performance of brown  
513 capuchins is likely to be comparable to those of chimpanzees in these AR discrimination  
514 experiments. It brings more evidence that monkeys can perform like apes in demanding cognitive  
515 tasks, for example like prerequisite ToM tests, contrary to previous scientists' common beliefs.  
516 These previous studies that found differences between performances of monkeys and apes in  
517 cognitive tasks, have normally compared apes that are highly habituated to experiments with naive  
518 monkeys (Amici et al., 2010). In addition to using similar standardized paradigm, our study tested  
519 highly habituated monkeys. Thereby, a comparison between our results and those on chimpanzees  
520 appears more reliable and appropriate.

521 Several recent studies on cognitive abilities comparing different monkey and ape species  
522 found no strong evidence of a difference between their performances (Amici et al., 2010; Meunier,  
523 2017; Schmitt et al., 2012, Tomasello et al., 1998). Instead, Amici and colleagues (2010) revealed a  
524 link in their results between cognitive capacities and social organisation: performances of species  
525 living in systems with fission-fusion dynamics (e.g. chimpanzees, bonobos, orangutans, and spider  
526 monkeys) exceeded those of species living in more stable groups (e.g. long-tailed macaques,  
527 gorillas and capuchin monkeys). Thereby, the type of social organisation seems to impact the  
528 physico-cognitive skills of primates (i.e. spatial memory, quantities, causality, etc.) and may  
529 provide better predictor of performance in some cognitive tasks than phylogenetic relatedness  
530 (Schmitt et al., 2012). This led to the development of the social intelligence hypothesis, which  
531 claims that primate intelligence evolved in response to challenges of living in large and complex  
532 groups (Byrne & Bates, 2010; MacLean et al., 2012; Schmitt et al., 2012). However, both for  
533 Tonkean macaques and brown capuchins, this explanation does not apply because they are not  
534 living in fission-fusion dynamics but instead in stable female-bonded groups with comparable social  
535 complexity, and are described as socially tolerant (Hare et al., 2003). These other social features  
536 could then interfere in their cognitive abilities' development. Thus, as socio-ecological conditions  
537 might better explain the presence of some cognitive abilities, instead of comparing monkeys and  
538 apes, we should focus on comparison between particular species. Studying with the same paradigm  
539 other nonhuman primates either closely related or presenting different social organisations (e.g.  
540 lemurs, white-faced capuchins, or long-tailed macaques) should provide insights about the origin of  
541 this ability and could help to clarify the social intelligence hypothesis (Byrne & Bates, 2010;  
542 Humphrey 1976; MacLean et al., 2012; Schmitt et al., 2012).

543       The positive results in both experiments could be explained by several alternative  
544 mechanisms of learning. First of all, an association learning between large or big food item and  
545 negative outcome could be a possibility. However, it seems to be an unlikely explanation because  
546 our successful subjects have chosen the item that appeared larger or bigger in the avoid tests, and  
547 they continued to choose the bigger or larger amount of items in each warm up preference trials  
548 before each session. Secondly, some might argue that successful individuals could have done  
549 reverse contingency learning, i.e. choose the smaller food item to obtain the larger one. However,  
550 this assumption is highly unlikely for the following reasons. First, the avoid control test clearly  
551 rules out this possibility. In fact, to validate this control, subjects needed to choose the apparent  
552 bigger or larger food item to obtain the truly bigger or larger one, and not the apparent smaller one  
553 as in previous tests. Second, several studies have already demonstrated that this type of contingency  
554 learning for both size and quantity of items is really difficult for these species, even for great apes,  
555 and require hundreds of trials to succeed in it (Anderson, Hattori, & Fujita, 2008; Boysen, Berntson,  
556 & Mukobi, 2001; Krachun et al., 2009; Vlamings, Uher, & Call, 2006).

557       In the Lens experiment, a few subjects did not succeed in different tests. First, a significant  
558 left pointing bias can explain the failure of two brown capuchins, one in the basic and the other one  
559 in the seen tracking test. The fact that this bias was not present in the initial conditions might  
560 indicate they do not understand the situation in the follow-up conditions and thus could also suggest  
561 that a more ecological procedure may have been necessary for these individuals for AR  
562 discrimination to be revealed. Second, Tonkean macaques and brown capuchins seem to have failed  
563 the Lens experiment in the last learning control because they have performed learning strategies  
564 based on different stimuli. Two Tonkean macaques and one brown capuchin obtained really low  
565 scores, demonstrating a significant preference for the truly smaller grape. This could be explained  
566 by a learning process to avoid the magnifying lens or to choose the grape that appeared with the size  
567 of the small one through the lenses. The other five unsuccessful Tonkean macaques and one brown  
568 capuchin obtained medium scores, which reflect a learning to choose the minimizing lens or the  
569 grape that appeared with the size of the bigger one through the lenses. Because neither the  
570 minimizing lens nor the grape appearing the size of the bigger one through lenses were no more  
571 available in this control, individuals could not anymore use these stimuli to succeed and so they  
572 chose randomly. Moreover, these individuals who failed the test still choose preferentially the  
573 bigger grape in warmup preference trials before each session, suggesting their failure was not due to  
574 a change of preference to smaller grapes.

575       In order to compare performances with those of chimpanzees, we carried out a second avoid-  
576 lens test at the end of the Lens experiment using exactly the same procedure as Krachun and  
577 colleagues in their study of 2016. Tonkean macaques' results obtained in the avoid-lens test n°2 are

quite inconsistent with those of avoid-lens test n°1. In fact, one individual succeeded in both but two individuals that failed avoid-lens test n°1 then succeeded in the avoid-lens test n°2, within one or two sessions only. For these two subjects, we carried out one more session with the unseen test procedure, in order to check whether by changing again sizes and lenses used, they will succeed (as before in the same unseen tracking test) and thus demonstrate a truly understanding of the illusion phenomena. However, both of them drastically failed by never choosing the truly larger grape in this session. These results highlight a learning strategy they must have developed during the four sessions of the avoid-lens test n°1. Because avoid-lens test n°2 was carried out after, individuals had the possibility to learn the new rule (larger grape placed behind magnifying lens). In addition, the only subject that passed the first avoid-lens control test needed four sessions to succeed in this last control. His result should be interpreted with caution because he could have also learned, faster than the others, i.e. within only three sessions. Surprisingly, the other Tonkean macaques and brown capuchins that have failed did not realise this potential quick learning: they both failed the avoid-lens n°1 and the avoid-lens n°2. A comparison between avoid-lens n°1 performances of Tonkean macaques and brown capuchins with avoid-lens test performances of chimpanzees in the study of Krachun and colleagues (2016) seems thus more reliable.

Instead of an absence of AR discrimination, individuals' failure could be interpreted by other explanations corresponding to methodological issues like too many experimenter's manipulations or isolation from their social group. Our test procedure could have required too much concentration, motivation and short-term memory for these animals, difficulties faced by Krachun and colleagues (2016) in their study on chimpanzees. Notably, poor performances of Tonkean macaques (poor success ratio, higher means of sessions to succeed in each test) compared to brown capuchins and chimpanzees may support this assumption: as we carried out half-sessions for brown capuchins, they had twice as many demo trials as Tonkean macaques and chimpanzees, hence more opportunities to understand the effects of stimuli illusion. Moreover, all brown capuchins came to the experimental room at least once per day, whereas with some Tonkean macaques, several days sometimes separated two successive sessions because they did not come back in the experimental room regularly. Because of those delays, Tonkean macaques might have faced memory challenges about the impact of illusion. Our results could thus underestimate their AR discrimination ability.

Given the pointing bias of some brown capuchins and the large failure of Tonkean macaques in the Lens experiment, ecological relevance of our paradigm has to be questioned. The better success ratio in the Mirror experiment could be due to the applied illusion. The mirror effect indeed creates a false presence of two raisins, which do not exist in reality. Whereas lens is modifying properties of grapes that already exist, it does not create new food items. Thereby, it is maybe easier to understand the effect of a mirror than the subtle modification of existent items created by lenses.

613 One alternative paradigm to test size illusion understanding could be the one developed by Karg  
614 and colleagues (2014). In their comparative study between great apes and children on AR  
615 discrimination, they created a size illusion using completion illusion, i.e. part of an object is  
616 partially occluding by another, in order to inverse size relation. As Tonkean macaques and brown  
617 capuchins are semi-arboreal primates, they might more regularly face size illusion challenges by  
618 completion illusion because of the dense vegetation in the canopy rather than by distortion.

619 Further research using different complementary types of illusion, e.g. colour or auditory  
620 illusions, could now bring more information about the scope of the AR understanding of these  
621 species. Another notable result is that several capuchins of various ages succeeded. Thus, AR  
622 discrimination ability seems to appear early in the development of brown capuchins. Further  
623 experiences on even younger individuals will be helpful to determine if the development of AR  
624 discrimination occurs around the same development stage as for human children and with the same  
625 developmental pattern (e.g. Hansen & Markman, 2005; Karg et al., 2014; Moll & Tomasello, 2012).

626 To conclude, our results support those of apes, demonstrating that this AR discrimination  
627 ability is not unique to great apes. As we have now demonstrated that even some monkey species  
628 seem to possess this understanding, it seems these abilities are perhaps more evolutionary ancestral  
629 than previously thought. More research on different monkey species with similar standardized  
630 procedure is now more than ever necessary to better understand evolution of our remarkable social  
631 cognition.

632

## 633 **Acknowledgments**

634 We are deeply grateful to all the members of the Centre de Primatologie de l'Université de  
635 Strasbourg for the welcome and very helpful technical assistance. Jamie Whitehouse is warmly  
636 thanked for insightful discussions and for editing the English of the manuscript. We also would like  
637 to sincerely thank the four anonymous reviewers and the editor for their constructive and very  
638 helpful suggestions.

639 This research did not receive any specific grant from funding agencies in the public,  
640 commercial, or not-for-profit sectors.

641

## 642 **References**

- 643 Amici, F., Aureli, F., & Call, J. (2010). Monkeys and Apes: Are Their Cognitive Skills Really so  
644 Different? *American Journal of Physical Anthropology*, 143(2), 188–197.
- 645 Amici, F., Visalberghi, E., & Call, J. (2009). Spider Monkeys (*Ateles Geoffroyi*) and Capuchin  
646 Monkeys (*Cebus Apella*) Follow Gaze around Barriers: Evidence for Perspective Taking?  
647 *Journal of Comparative Psychology*, 123(4), 368–374.
- 648 Anderson, J. R., Hattori, Y., & Fujita, K. (2008). Quality before Quantity: Rapid Learning of  
649 Reverse- Reward Contingency by Capuchin Monkeys (*Cebus Apella*). *Journal of Comparative*

650 *Psychology*, 122(4), 445–448.

651 Baillargeon, R., Scott, R. M., & He, Z. (2010). False-Belief Understanding in Infants. *Trends in*

652 *Cognitive Sciences*, 14(3), 110–118.

653 Baron-Cohen, S. (1989). Are Autistic Children “Behaviorists”? An Examination of Their Mental-

654 Physical and Appearance-Reality Distinctions. *Journal of Autism and Developmental Disorders*,

655 19(4), 579–600.

656 Boysen, S. T., Berntson, G. G., & Mukobi, K. L. (2001). Size Matters: Impact of Item Size and

657 Quantity on Array Choice by Chimpanzees (Pan Troglodytes). *Journal of Comparative*

658 *Psychology*, 115(1), 106.

659 Braine, M., & Shanks, B. L. (1965). The Development of Conservation of Size. *Journal of Verbal*

660 *Learning and Verbal Behavior*, 4(3), 227–242.

661 Bräuer, J., Call, J., & Tomasello, M. (2007). Chimpanzees Really Know What Others Can See in a

662 Competitive Situation. *Animal Cognition*, 10(4), 439–448.

663 Bray, J., Krupenye, C., & Hare, B. (2014). Ring-Tailed Lemurs (*Lemur Catta*) Exploit Information

664 about What Others Can See but Not What They Can Hear. *Animal Cognition*, 17(3), 735–744.

665 Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2007). Enculturated Chimpanzees Imitate

666 Rationally. *Developmental Science*, 10(4), F31–F38.

667 Byrne, R. W., & Bates, L. A. (2010). Primate Social Cognition: Uniquely Primate, Uniquely Social,

668 or Just Unique? *Neuron*, 65(6), 815–830.

669 Call, J., & Tomasello, M. (2008). Does the Chimpanzee Have a Theory of Mind? 30 Years Later.

670 *Trends in Cognitive Sciences*, 12(5), 187–192.

671 Canteloup, C., Bovet, D., & Meunier, H. (2015). Do Tonkean Macaques (*Macaca Tonkeana*) Tailor

672 Their Gestural and Visual Signals to Fit the Attentional States of a Human Partner? *Animal*

673 *Cognition*, 18(2), 451–461.

674 Canteloup, C., Piraux, E., Poulin, N., & Meunier, H. (2016). Do Tonkean Macaques (*Macaca*

675 *Tonkeana*) Perceive What Conspecifics Do and Do Not See? *PeerJ*, 4:e1693.

676 Defolie, C., Malassis, R., Serre, M., & Meunier, H. (2015). Tufted Capuchins (*Cebus Apella*) Adapt

677 Their Communicative Behaviour to Human’s Attentional States. *Animal Cognition*, 18(3), 747–

678 755.

679 Flavell, J. H. (1993). The Development of Children’s Understanding of False Belief and the

680 Appearance-Reality Distinction. *International Journal of Psychology*, 28(5), 595–604.

681 Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the Appearance-Reality

682 Distinction. *Cognitive Psychology*, 15, 95–120.

683 Flavell, J. H., Green, F. L., Flavell, E. R., Watson, M. W., & Campione, J. C. (1986). Development

684 of Knowledge about the Appearance-Reality Distinction. *Monographs of the Society for*

685 *Research in Child Development*, 51(1): i.

686 Flavell, J. H., Green, F. L., Wahl, K. E., & Flavell, E. R. (1987). The Effects of Question

687 Clarification and Memory Aids on Young Children’s Performance on Appearance-Reality

688 Tasks. *Cognitive Development*, 2(2), 127–144.

689 Flombaum, J. I., & Santos, L. R. (2005). Rhesus Monkeys Attribute Perceptions to Others. *Current*

690 *Biology*, 15(5), 447–452.

691 Gopnik, A., & Astington, J. W. (1988). Children’s Understanding of Representational Change and

692 Its Relation to the Understanding of False Belief and the Appearance-Reality Distinction. *Child*

693 *Development*, 59(1), 26.

694 Hansen, M. B., & Markman, E. M. (2005). Appearance Questions Can Be Misleading: A

695 Discourse-Based Account of the Appearance–Reality Problem. *Cognitive Psychology*, 50(3),

696 233–263.

697 Hare, B., Addessi, E., Call, J., Tomasello, M., & Visalberghi E. (2003). Do Capuchin Monkeys,

698 *Cebus Apella*, Know What Conspecifics Do and Do Not See? *Animal Behaviour*, 65(1), 131–

699 142.

700 Hare, B., Call, J., & Tomasello, M. (2001). Do Chimpanzees Know What Conspecifics Know?

701 *Animal Behaviour*, 61(1), 139–151.

- 702 Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees Deceive a Human Competitor by Hiding.  
703 *Cognition*, 101(3), 495–514.
- 704 Hirata, S., & Matsuzawa, T. (2001). Tactics to Obtain a Hidden Food Item in Chimpanzee Pairs  
705 (*Pan Troglodytes*). *Animal Cognition*, 4(3–4), 285–295.
- 706 Humphrey, N.K. 1976. in *Growing Points in Ethology*, P. P. G. Bateson, R. A. Hinde, Eds.  
707 *Cambridge Univ. Press, Cambridge*, pp. 303–321.
- 708 Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2014). All Great Ape Species (*Gorilla Gorilla*,  
709 *Pan Paniscus*, *Pan Troglodytes*, *Pongo Abellii*) and Two-and-a-Half-Year-Old Children (*Homo*  
710 *Sapiens*) Discriminate Appearance from Reality. *Journal of Comparative Psychology*, 128(4),  
711 431–439.
- 712 Kelley, L. A., & Kelley, J. L. (2014). Animal Visual Illusion and Confusion: The Importance of a  
713 Perceptual Perspective. *Behavioral Ecology*, 25(3), 450–463.
- 714 Krachun, C., Call, J., & Tomasello, M. (2009). Can Chimpanzees (*Pan Troglodytes*) Discriminate  
715 Appearance from Reality? *Cognition*, 112(3), 435–450.
- 716 Krachun, C., Lurz, R., Russell, J. L., & Hopkins, W. D. (2016). Smoke and Mirrors: Testing the  
717 Scope of Chimpanzees' Appearance–Reality Understanding. *Cognition*, 150, 53–67.
- 718 Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that  
719 other individuals will act according to false beliefs. *Science*, 354(6308), 110–114.
- 720 MacLean, E. L., Matthews, L. J., Hare, B. A., Nunn, C.L., Anderson, R.C., Aureli, F., ... Wobber,  
721 V. (2012). How Does Cognition Evolve? Phylogenetic Comparative Psychology. *Animal*  
722 *Cognition*, 15(2), 223–238.
- 723 Martin, A., & Santos, L. R. (2016). What Cognitive Representations Support Primate Theory of  
724 Mind? *Trends in Cognitive Sciences*, 20(5), 375–382.
- 725 Matsuno, T., & Fujita, K. (2009). A Comparative Psychophysical Approach to Visual Perception in  
726 Primates. *Primates*, 50(2), 121–130.
- 727 Meunier, H. (2017). Do Monkeys Have a Theory of Mind? How to Answer the Question?  
728 *Neuroscience & Biobehavioral Reviews*, 82, 110–123.
- 729 Moll, H., & Tomasello, M. (2012). Three-Year-Olds Understand Appearance and Reality—Just Not  
730 about the Same Object at the Same Time. *Developmental Psychology*, 48(4), 1124–1132.
- 731 Overduin-de Vries, A. M., Spruijt, B. M., & Sterck, E. H. M. (2014). Long-Tailed Macaques  
732 (*Macaca Fascicularis*) Understand What Conspecifics Can See in a Competitive Situation.  
733 *Animal Cognition*, 17(1), 77–84.
- 734 Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The*  
735 *Behavioural & Brain Sciences*, 4, 515–526.
- 736 Quinn, G.P., & Keough, M.J. 2002. Experimental designs and data analysis for biologists.  
737 *Cambridge University Press. Cambridge*.
- 738 Sapp, F., Kang L., & Muir, D. (2000). Three-Year-Olds' Difficulty with the Appearance–Reality  
739 Distinction: Is It Real or Is It Apparent? *Developmental Psychology*, 36(5), 547–560.
- 740 Schmitt, V., Pankau, B., & Fischer, J. (2012). Old World Monkeys Compare to Apes in the Primate  
741 Cognition Test Battery. *PLoS ONE*, 7(4), e32024.
- 742 Tomasello, M., Call, J., & Hare, B. (1998). Five Primate Species Follow the Visual Gaze of  
743 Conspecifics. *Animal Behaviour*, 55(4), 1063–1069.
- 744 Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees Understand Psychological States – the  
745 Question Is Which Ones and to What Extent. *Trends in Cognitive Sciences*, 7(4), 153–156.
- 746 Visalberghi, E., & McGrew, W. C. (1997). Cebus meets Pan. *International Journal of Primatology*,  
747 18(5), 677–681.
- 748 Vlamings, P. H., Uher, J., & Call, J. (2006). How the Great Apes (*Pan Troglodytes*, *Pongo*  
749 *Pygmaeus*, *Pan Paniscus*, and *Gorilla Gorilla*) Perform on the Reversed Contingency Task: The  
750 Effects of Food Quantity and Food Visibility. *Journal of Experimental Psychology: Animal*  
751 *Behavior Processes*, 32(1), 60–70.
- 752 Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development:



753       The Truth about False Belief. *Child Development*, 72(3), 655–684.  
754       Whiten, A., & Byrne, R. W. (1988). Tactical Deception in Primates. *Behavioral and Brain*  
755       *Sciences*, 11(02), 233.  
756       Woodruff, G., & Premack, D. (1979). Intentional Communication in the Chimpanzee: The  
757       Development of Deception. *Cognition*, 7(4), 333–362.  
758  
759  
760