

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

Testing the relationship between looking time and choice preference in long-tailed macaques

V. Wilson^{1,2,3,*}, C. Kade¹, J. Fischer^{1,3,4}

¹Cognitive Ethology Laboratory, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany

²Department of Primate Cognition, Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology, University of Göttingen, Göttingen, Germany

³Leibniz ScienceCampus Primate Cognition

*Corresponding author: vanessa.wilson@unine.ch

29 **Abstract**

30 Visual bias in social cognition studies is often interpreted to indicate preference, yet it is
31 difficult to elucidate whether this translates to social preference. Moreover, visual bias is
32 often framed in terms of surprise or recognition. It is thus important to examine whether an
33 interpretation of preference is warranted in looking time studies. Here, using touchscreen
34 training, we examined (1) looking time to non-social images in an image viewing task, and
35 (2) preference of non-social images in a paired choice task, in captive long-tailed macaques,
36 *Macaca fascicularis*. In a touchscreen test phase, we examined (3) looking time to social
37 images in a viewing task, and (4) preference of social images in a paired choice task. Finally,
38 we examined (5) looking time to social images in a non-test environment. For social content,
39 the monkeys did not exhibit clear preferences for any category (conspecific/heterospecific, in-
40 group/outgroup, kin/non-kin, young/old) in the explicit choice paradigm, nor did they
41 differentiate between images in the viewing tasks, thus hampering our interpretation of the
42 data. Post-hoc analysis of the training data however revealed a visual bias towards images of
43 food and objects over landscapes in the viewing task. Similarly, across choice-task training
44 sessions, food and object images were chosen more frequently than landscapes. This suggests
45 that the monkeys' gaze may indeed indicate preference, but this only became apparent for
46 non-social stimuli. Why these monkeys had no biases in the social domain remains enigmatic.
47 To better answer questions about attention to social stimuli, we encourage future research to
48 examine behavioral measures alongside looking time.

49

50 **Keywords:** *Macaca fascicularis*, visual bias, social cognition, attention, gaze

51

52

53

54 **Introduction**

55 In social cognition research, differential allocation of visual attention (i.e. looking time) has
56 become a popular way to measure response to stimuli (Winters, Dubuc, & Higham, 2015). As
57 an implicit measure, looking time provides a simple way to test responses to stimuli with
58 minimum training. Originally developed for use with human infants (Fantz, 1958), looking
59 time paradigms have since been adopted across a broad range of species, including corvids,
60 canines and cichlids, but especially amongst primates (Albuquerque et al., 2016; Bird &
61 Emery, 2008; Dufour, Pascalis, & Petit, 2006; Kohda et al., 2015; Rakotonirina, Kappeler, &
62 Fichtel, 2018). Such studies have used this paradigm to assess attentional differences to social
63 information, such as group membership (Méary et al., 2014), familiarity (Adams &
64 Macdonald, 2018), age (Sato et al., 2012), sex (Koba & Izumi, 2008), facial coloration (Waitt
65 & Little, 2006), gaze direction (Farroni, Menon, & Johnson, 2006), trustworthiness (Costa et
66 al., 2018), emotional valence (Bethell, Holmes, MacLarnon, & Semple, 2012), symmetry
67 (Paukner et al., 2017), kinship (Pfefferle et al., 2014) and social bonds (Almeling et al., 2016).

68 Assessing allocation of social attention is important in understanding what variables
69 play a role in social decision making. For example, in rhesus macaques (*Macaca mulatta*),
70 male facial skin coloration has been proposed as a sexually selected trait that indicates fitness.
71 In line with this theory, female rhesus macaques exhibit longer looking time to images of
72 male faces with a redder coloration (Waitt et al., 2003), and it thus follows that, in a free-
73 ranging colony, females are more likely to solicit males with dark red over pale faces (Dubuc,
74 Allen, Maestripieri, & Higham, 2014). Looking time paradigms, thus placed in the context of
75 social behavior, can be informative about social decisions, such as mate choice.

76 A limitation to looking time paradigms, however, is their interpretation (Winters et al.,
77 2015). An initial issue is the use of the term ‘visual preference’ to describe prolonged
78 attention to a stimulus. The use of this term is problematic as it often ambiguously suggests

79 that a visual preference could translate to a social preference, an implication that is
80 unsubstantiated without additional behavioral data. Winters et al. (2015) suggest the use of
81 ‘visual bias’ as a less misleading term, which simply implies discrimination between two or
82 more stimuli. A secondary problem of interpreting looking time is in understanding why there
83 is a visual bias, an issue that has also been raised in the field of infant cognition. As with
84 nonhuman subjects, looking time paradigms have provided a window into how infants
85 perceive and discriminate stimuli. However, Aslin (2007) noted the distinct lack of clear
86 linking hypotheses for such discrimination studies, that is, what should looking time tell us
87 about the underlying cognitive process in question? Framing this in primate terms, a number
88 of studies have examined how study subjects respond to images of conspecifics versus
89 heterospecifics. Sumatran orangutans (*Pongo abelii*), humans (*Homo sapiens*), brown
90 capuchins (*Sapajus apella*), rhesus, Tonkean, Japanese, bonnet and pig-tailed macaques
91 (*Macaca mulatta*, *M. tonkeana*, *M. fuscata*, *M. radiata*, *M. nemestrina*) and red-fronted
92 lemurs (*Eulemur rufifrons*) look longer at images of conspecifics versus heterospecifics
93 (Adams & Macdonald, 2018; Dufour et al., 2006; Fujita, 1987; Rakotonirina et al., 2018).
94 Attending to conspecific stimuli could be advantageous as they contain socially relevant cues
95 (Rakotonirina et al., 2018). Yet contrastingly, studies in rhesus and stump-tailed macaques
96 (*M. arctoides*) report visual bias for heterospecific over conspecific images (Fujita, 1987;
97 Méary et al., 2014). In this case, it is difficult to assert that subjects exhibit a preference for a
98 certain stimulus type, as the underlying motivation for visual bias is unclear. That is, do
99 subjects look longer due to socially relevant cues, or because they see something novel that
100 captures their interest? Such responses are perhaps better framed in terms of interest rather
101 than preference, at least until more is understood about what drives a visual bias within this
102 context.

103 There is an additional, paradoxical, problem in the looking time literature: attention
104 has been interpreted to indicate anticipation or expectation. Anticipatory gaze was used by
105 Krupenye et al. (2016) in examining apes' understanding of false belief. A scenario depicted
106 an observer looking for a displaced object. Test participants looked more often towards the
107 place where the observer last saw the object than where the object was last placed, suggesting
108 that they anticipated where the observer would look for the object. A similar interpretation
109 was made in assessing captive chimpanzees' responses to visual depictions of infanticide.
110 Despite having experienced infanticide, chimpanzees still looked longer at infanticide scenes
111 over conflict scenes, suggesting that infanticide violated their expectations of regular social
112 behavior (to care for an infant; Rudolf Von Rohr et al., 2015). Similarly, male long-tailed
113 macaques looked longer at images containing incongruent information about group
114 dominance hierarchy, such as a dominant individual being submissive to a subordinate
115 (Overduin-de Vries et al., 2016). In contrast, chimpanzees did not look longer at scenes
116 depicting unexpected social outcomes, instead attending primarily to negatively valenced
117 emotional content (Wilson et al., in prep). The point here is that depending on the paradigm in
118 which a stimulus is framed, the interpretation of attention varies from one of preference,
119 recognition or interest to one of expectation or violation of expectation. The problem is that,
120 regardless of the paradigm presented, underlying characteristics that motivate attention still
121 vary (Winters et al., 2015), an issue that seems to be ubiquitous across species, including
122 humans. Moreover, reasons for why one stimulus might motivate a longer looking time than
123 another stimulus could vary within a study, for example, gaze allocation might depend on a
124 combination of or alternation between factors such as interest, recognition, preference and
125 surprise. The potential multi-faceted nature of looking time is therefore worthy of more
126 detailed consideration.

127 Given the extent to which gaze responses are used in cognitive research, it is important
128 to address underlying motivations of visual bias and thereby clarify interpretations regarding
129 looking time. Returning to the concept of preference, there is some evidence that visual bias
130 may in fact play a role in explicit choices. For example, when human subjects are
131 manipulated to look at a stimulus for longer, then they are more likely to select that stimulus
132 than a counterpart that is presented for a shorter period (Shimojo et al., 2003). Similarly, in a
133 forced-choice paradigm, women with a preference for masculine faces exhibited stronger gaze
134 patterns to preferred stimuli (Lyons et al., 2016), suggesting that visual bias may reflect
135 preferences. The limitation here is that gaze and preference were assessed in the same task,
136 and thus could not be decoupled. In the current study, we tested the idea that visual bias
137 indicates choice preference, by comparing looking time in a viewing task to explicit choice in
138 a separate preference task, in a group of captive long-tailed macaques.

139 We first trained the monkeys on a novel viewing task, where they could control the
140 time the image was presented, and a separate choice task, where they had to choose between
141 two images. For training, we used images of food, landscapes and novel objects, with a
142 separate set of images for each task to maintain novelty. To assess for preferences during
143 training on the choice task, we predicted that if the monkeys perceived the images as
144 meaningful representations (Fagot, 2000; Fagot et al., 2010), they would choose images of
145 food over landscape or objects, and images of objects over landscape, as both food and novel
146 objects (e.g. tools, keys, glasses) motivate their interest (see Appendix for more details). We
147 carried out unplanned analyses of the training data. Due to the use of different image sets
148 between tasks, we could not, however, examine the relationship between looking time and
149 choice.

150 To test the monkeys' looking time and preferences for social images, we presented
151 them with a different viewing task where images were presented for a fixed time, as well as a

152 choice task that used the same paradigm as for the non-social stimuli. We presented four types
153 of social image pairings: kin vs. nonkin, young vs. old, ingroup vs. outgroup, and conspecifics
154 vs. heterospecifics. The same images were used in both tasks. We expected longer looking
155 times for outgroup over ingroup (Gothard et al., 2004; Schell et al., 2011), younger over older
156 monkeys (Almeling et al., 2016) and non-kin over kin (Pfefferle et al., 2014). We made no
157 directional predictions for species, given previous conflicting findings (Adams & Macdonald,
158 2018; Dufour et al., 2006; Fujita, 1987; Méary et al., 2014; Rakotonirina et al., 2018). We
159 expected that if monkeys' looking time reflected preference, then we would see a relationship
160 between looking time and choice across image categories. Additionally, we conducted a
161 separate 'free view' task in the monkeys' home enclosure, allowing us to validate our findings
162 for the looking task results without reward motivation and in a bigger sample. For this task we
163 used only the age category images and thus our predictions were consistent with those for the
164 previous looking task.

165

166 **Methods and Results**

167 We collected data from a group of 36 long-tailed macaques (*Macaca fascicularis*) housed at
168 the German Primate Center, Göttingen. The group had access to both an indoor (49m²) and
169 outdoor (141m²) enclosure, with ad libitum access to food, water and enrichment. Details of
170 participating monkeys can be found in Table 1 of the Appendix.

171 We collected data in three experimental phases, on five different tasks. Experiment I
172 consisted of training monkeys on two different touchscreen tasks (viewing task I and choice
173 task I) using non-social images. Experiment II consisted of testing monkeys' responses to
174 social images on two different touchscreen tasks (viewing task II and choice task II).
175 Experiment III consisted of collecting data on the free view task, where monkeys viewed
176 images presented in their home enclosure. In all three viewing tasks, images were presented

177 individually, and we measured looking time to each image, defined by the monkeys clearly
178 orienting their head and eyes towards the images. The choice tasks presented the monkeys
179 with pairs of images. In choice task I, they had to learn to select one image to progress to the
180 next trial and subsequently, the reward. We then tested their social preferences on this
181 paradigm in choice task II. For this we measured type, and side, of image chosen. All tasks
182 except the free view task took place in test cubicles where monkeys interacted with a touch
183 screen. The free view task took place in the monkeys' enclosure (see Table 1 of the
184 Appendix). Details of each task are provided below.

185

186 ----- Figure 1 here -----

187

188 **Testing apparatus.** Individuals participated voluntarily in all touch screen testing,
189 which took place in a separated testing area divided into six cubicles (height: 2.6 m; width:
190 2.25 m; depth: 1.25 m). Monkeys were rewarded for participation during testing with diluted
191 grape juice (diluted at 200ml juice:100ml water; 0.25 ml per reward). Testing hours were
192 Monday to Friday, 10:00 – 12:00 and 14:00 – 18:00.

193 For the touch screen tasks, monkeys were trained and tested individually. All monkeys
194 received prior training on the touch screen before commencing any task. When they exhibited
195 excess caution towards the screen, for example by retreating when a stimulus appeared on-
196 screen, we habituated them by pairing them with a partner who was comfortable interacting
197 with the touch screen display. We used the XBI touch screen – eXperimental Behavioural
198 Instrument (Calapai et al., 2016), with a 15" Elo touchmonitor, MacBook Air (macOS Sierra
199 Version 10.12.6), HD wifi camera with wide angle lens, and two pumps and drinking tube for
200 automated fluid reward (see Figure 1). Experiments were run using MWorks (version 0.7).

201 **Video coding.** Looking time (in viewing task I, viewing task II and the free view task)
202 as well as which side of the screen the monkey touched first, and which image they touched
203 first (choice tasks I and II) were coded from videos (image chosen was recorded by MWorks,
204 but sometimes the monkeys' first touch to an image did not activate it, so we checked for this
205 when coding the videos). We used the free behavioral coding software SolomonCoder
206 (version 17.03.22). For each task (excluding choice task I; see Appendix for details), a subset
207 of videos was coded by two coders to assess interrater reliability. At least one of the coders
208 was always naïve to the hypotheses.

209 **General analytical procedure.** Analyses were conducted in R studio version 1.0.153. We ran
210 a series of mixed effects models using the lme4 package. For all models, we ran bootstrapped
211 95% confidence intervals. Where we adjusted models with a Bonferroni correction, we report
212 the corresponding adjusted confidence intervals. For the viewing tasks, our analyses excluded
213 images the monkeys did not look at. For viewing task I, choice task I and II, and visual bias to
214 image category in viewing task II, we set the number of adaptive Gauss-Hermite quadrature
215 points to zero to aid model convergence. The number of Gauss-Hermite quadrature points
216 (nAGQ) determines how the random effects are distributed in the model. When fitting glms
217 using the lme4 package in R, nAGQ defaults to one (Bolker, 2019; Zhang et al., 2011).
218 Setting nAGQ to zero means that parameter estimation is less exact, but faster (Bolker, 2019),
219 since the computational time decreases with fewer quadrature points (Kim al., 2013). Each
220 model is described in detail below.

221 **Experiment I.**

222 Experiment I involved training monkeys on two tasks displaying images of non-social
223 content.

224 **Viewing task I.** We examined looking time to images in the first training session only,
225 when the images were novel (see Appendix for details about further sessions).

226 *Participants.* We report data from thirteen monkeys trained on this task (6M; 7F; M
227 age = 5.91 years, $SD = 5.47$), eight of which received their initial training on a different
228 apparatus (see Appendix, Table 1 and methods, for details).

229 *Stimuli.* Sixty images of non-social content were used. Images were divided into three
230 categories - landscape, object and food - with equal pairings between each category. Images
231 were scaled in GIMP to 4000 x 3000 pixels at 72 ppi. For more details see the Appendix.

232 *Procedure.* An image appeared on screen (19.63cm x 19.63cm) with a target beneath
233 (grey square: 3.93cm x 3.93cm; see Figure 1). The monkeys could view the image for 60 s,
234 or, by pressing the target, change the image sooner. The purpose of this was to give the
235 monkeys control over how long they viewed each image. Each session consisted of 20
236 images. Reward was given at random intervals, so monkeys were not reinforced to touch the
237 target but learned to change the picture by trial and error, thereby removing food-based
238 incentives for image viewing. Order of image presentation was randomized. See the Appendix
239 for additional details.

240 *Interrater reliability.* We calculated interrater reliability for looking time to each
241 image for seven monkeys, using intraclass correlation coefficients (Koo & Li, 2016; Shrout &
242 Fleiss, 1979). We estimated the reliability of mean values across k coders, $ICC(3,k)$. For
243 viewing task I we report an $ICC(3,k)$ of 0.72, indicating good reliability (Koo & Li, 2016).

244 *Models.* To examine differences in initial looking time to each non-social image
245 category, we ran three pair-wise generalized linear mixed models, fitted with a Gamma
246 distribution, adjusted with Bonferroni correction. The models contrasted 1) food versus
247 landscape, 2) landscape versus object, and 3) food versus object. Looking time was the
248 dependent variable, image category was the binomial fixed effect and ID was a random effect.
249 Looking time to images of food and objects was significantly longer than for landscapes, even

250 after adjusting for multiple comparisons. There was no significant difference in looking time
251 between objects and food (see Table 1 and Figure 2, left).

252

253 ----- Table 1 here -----

254

255 ***Choice task I.***

256 *Participants.* Thirteen monkeys were trained on this task, eleven of which took part in
257 viewing task I (6M, 7F; *M* age = 5.79 years, *SD* = 5.58; number training sessions: *M* = 18.71,
258 *SD* = 9.23, range = 6-32 sessions).

259 *Stimuli.* Stimuli are as described for viewing task I, except we used separate training
260 sets for each task to ensure that images were novel and therefore promoted interest in the task.

261 *Procedure.* We presented monkeys with pairs of images from which they could make
262 a choice. Each trial started with a central target (green circle, 3.93cm x 3.93cm) on the screen,
263 which provided reward when touched. Next, two images were presented on either side of the
264 screen (7.85cm x 7.85cm). Order and side of image presentation were randomized.

265 Participants had to select one image, though they were not rewarded for this – by providing
266 juice only at the start of each trial (upon touching the target circle), reward was decoupled
267 from image viewing. Image pairs were presented for up to 300 seconds, after which a new
268 trial would start. When an image was selected, the other image would disappear and the
269 selected image would increase in size (15.7cm x 15.7cm; see Figure 1) and remain for six
270 seconds, before the next trial started.

271 Each session consisted of 20 trials each. Monkeys received a minimum of six training
272 sessions, and reached criterion when they selected an image on every trial, within 20 seconds,
273 over two consecutive sessions.

274 *Models.* We examined image choice across training sessions for seven monkeys that
275 passed training, using a generalized linear mixed Poisson model. We excluded any incomplete
276 training sessions ($n = 21$ for five monkeys). Number of instances of each category chosen per
277 monkey/session was the dependent variable, image category was the fixed effect and ID was
278 the random effect. In a full model, we compared food and object choices over landscape.
279 Consistent with looking time, monkeys chose food and object images more than landscape. A
280 subset model that compared only food and object choices revealed that monkeys chose food
281 images significantly more than object images (see Table 2 and Figure 2, right). Due to some
282 procedural adjustments during training (see Appendix for details), we ran additional models
283 to determine whether these changes had any overall effect on choice preferences. Results
284 were consistent with our full model, reflecting preferences for food and object over landscape
285 in both early and later trials (see Appendix, Results and Figure 1, for details).

286

287 ----- Table 2 here -----

288

289 To check for side bias of image selection, we included which side image chosen as the
290 binary dependent variable, image category chosen as the fixed effect and ID as a random
291 effect. We found no indication that which side image was chosen (left or right of screen) was
292 related to image category chosen (food versus landscape: $OR = 0.94$, $SE = 0.12$, 95% $CI =$
293 $[0.73, 1.18]$; object versus landscape: $OR = 1.03$, $SE = 0.12$, 95% $CI = [0.80, 1.29]$).

294 As the images used for viewing task I and choice task I were, for the purposes of
295 training, different, we could not directly examine the relationship between looking time and
296 choice for these two tasks. The results do, however, indicate that preferences for images of
297 food and objects in choice task I were consistent with longer looking times to novel food and

298 object images in viewing task I. This suggests that there may be a relationship between
299 looking time and category preference for non-social images.

300

301 ----- Figure 2 here -----

302

303 **Experiment II.**

304 Experiment II involved testing monkeys on two tasks displaying images of social content.

305 *Viewing task II.*

306 Viewing task II was developed as an alternative testing procedure to that developed
307 with viewing task I. Although six monkeys reached criterion on viewing task I, testing using
308 this method was not feasible (see Appendix for details). Using data from viewing task II, we
309 examined the relationship between looking time to social images, and selection of social
310 images in Choice task II. Choice task II and viewing task II were counterbalanced in order of
311 presentation.

312 *Participants.* Eleven monkeys were tested, but one monkey did not continue to engage
313 the apparatus and thus failed to complete the test, so data were analyzed from ten monkeys,
314 seven of which took part in viewing task I (6M, 4F; M age = 5.07, SD = 3.96).

315 *Test stimuli.* Test stimuli were the same images of social content as for the choice task.

316 *Procedure.* The monkeys started each trial by pressing a target (purple diamond;
317 3.93cm x 3.93cm) to receive a reward. They were then presented with one image (15.7cm x
318 15.7cm; Figure 1) for six seconds, randomized in order. As the monkeys were familiar with
319 the target-reward system, no training was required. Each monkey received one test session
320 consisting of 36-46 trials, which varied due to number of kin images available per subject
321 (between 18 and 23 image pairs: M = 32.3 trials, SD = 5.89).

322 *Interrater reliability.* Viewing times were coded for interrater reliability from five
323 videos. ICC(3,*k*), was 0.92, indicating excellent reliability (Koo & Li, 2016).

324 *Models.* To examine visual bias towards social images, we ran separate generalized
325 linear mixed models for each image category, kin, age, group and species. Viewing time was
326 the dependent variable, type of image viewed was the binomial fixed effect, and ID the
327 random effect. Looking time did not differ significantly between image type within any of the
328 four social image categories (see Table 3).

329 Finally, we examined whether looking time to images decreased across trials. We ran
330 a linear mixed model with log transformed looking time as the dependent variable, trial
331 number as a fixed effect and ID as a random effect. Looking time did not decrease
332 significantly across trials ($b = 0.003$, $SE = 0.003$, 95% CI = [-0.004, 0.01]).

333 To account for some variation in image content, we additionally conducted sensitivity
334 analyses, finding no effect of variation in image content on looking time (see Appendix).

335

336 ***Choice task II.***

337 This task built on training from choice task I and used the same paradigm.

338 *Participants.* Seven monkeys that reached criterion in choice task I were tested on this
339 task (5M; 2F; M age = 5.89 years, $SD = 4.46$). All seven monkeys also took part in viewing
340 task II.

341 *Test stimuli.* Monkeys were presented with images of kin vs. nonkin (one-six pairs),
342 ingroup vs. outgroup (four pairs), conspecifics vs. heterospecifics (four pairs) and young vs.
343 old (eight-nine pairs). Images were sorted into pairs and matched by sex, approximate age and
344 rank, and gaze direction. Images were cropped and scaled in GIMP to dimensions of 3000 x
345 3000 pixels at 72 ppi.

346 *Procedure.* Each monkey received one test session. Session length consisted of 18 to
347 23 trials which varied due to number of kin images available per subject ($M = 20.71$; $SD =$
348 1.89).

349 *Interrater reliability.* From six videos, reliability of side chosen and first touch were
350 assessed by checking for discrepancies between coders. Three discrepancies for side chosen
351 were found, only one of which indicated an error by the primary coder. This was corrected in
352 data analysis. Two discrepancies were found for first touch, neither of which were errors by
353 the primary coder.

354 *Models.* There were no clear group-level preferences for image choice. To examine
355 whether side chosen predicted image choice, we ran four generalized linear mixed models,
356 one per image category, with side chosen as the dependent binomial variable, image type
357 chosen as the fixed effect and ID as the random effect. Which side of the screen image was
358 chosen was not related to the type of image chosen (age: $OR = 2.34$, $SE = 0.64$, 95% $CI =$
359 $[0.68, 9.30]$; group: $OR = 1.17$, $SE = 1.07$, 95% $CI = [0.03, 21.35]$; kin: $OR = 0.56$, $SE = 1.10$,
360 95% $CI = [< .001, 7.00]$; species: $OR = 2.10$, $SE = 0.87$, 95% $CI = [0.39, 26.67]$).

361 To examine whether looking time in viewing task II predicted image choice in the
362 choice task, we aggregated looking time to each image group per monkey, and paired this
363 with the number of instances images in each group were chosen. We ran a generalized linear
364 mixed model for each image category, with a Gamma distribution, looking time as the
365 dependent variable, instances images were chosen as fixed effects and ID as a random effect.
366 Image choice did not predict looking time to images in viewing task II across image
367 categories (see Table 3, grey rows).

368 To examine whether there was any effect of test order (i.e. order in which monkeys
369 received the choice task II or viewing task II) on response in the viewing task, we ran a
370 generalized linear mixed model with a Gamma distribution, with looking time as the

371 dependent variable, order as a fixed effect and ID as a random effect. Seeing the images in the
372 choice task first did not have any effect on looking time to the images when receiving the
373 viewing task second ($b = 0.43$, $SE = 0.28$, 95% CI = [-0.47, 1.24]).

374

375 **Experiment III.**

376 Experiment III involved testing monkeys without the use of the touchscreen in their home
377 enclosure.

378 *Free view task.*

379 This task was conducted in the monkeys' enclosure, to verify that our touch screen results
380 were not simply an artifact of the testing environment. That is, we carried out the test without
381 providing food reward as a motivator, and were also able to test a larger sample by including
382 monkeys that do not usually come into the testing cubicles. We followed the methods of
383 Schell et al. (2011) and Almeling et al. (2016).

384 *Participants.* We tested 29 monkeys (10M, 19F; M age = 7.22 years, $SD = 6.82$), 14 of
385 which had participated in at least one of the touchscreen tasks.

386 *Test stimuli and apparatus.* Following results for Almeling et al., (2016), we chose
387 two images from the age category of the prior touch screen studies, depicting the same female
388 at two different ages. We selected images for a female who was not a regular test subject, as
389 we were not sure if we would be able to test all non-regular test subjects; we also wished to
390 re-test our regular test subjects in their home enclosure to examine whether their responses to
391 photos were consistent with responses in the testing cubicles. Each image was cropped to a
392 circular frame with a 17 cm diameter, printed on A4 paper with a white background and set
393 inside a wooden frame (Figure 1). Each subject saw both conditions. Condition order (i.e.
394 young vs. old) was semi-randomly counterbalanced.

395 *Procedure.* During testing, which image to present was determined by an assistant
396 from a pre-determined order of images to be shown to each subject. The experimenter was
397 blind to the conditions. Subjects were tested alone in one area of the enclosure and sitting near
398 to the mesh. The experimenter would approach within 1 m, holding the frame at eye level to
399 the monkey with a handheld camera (Panasonic HC-X909 and HC-X929, full HD with wide
400 angle lens) above the frame. The experimenter wore sunglasses and kept their head tilted
401 down to avoid eye contact with the monkey. The image was covered with a white sheet prior
402 to testing. The experimenter would remove this at the start of the experiment and the assistant
403 would start a timer. The image was kept in place for 60 s, or less if the monkey left, or
404 another monkey approached. Monkeys were not rewarded in this task.

405 In four instances we repeated a condition (three times for the ‘old’ image condition,
406 one time for the ‘young’ image condition): twice because the experiment was disturbed by
407 another monkey in the first three seconds, and twice because the test subject was distracted
408 and left almost immediately.

409 *Interrater reliability.* Viewing time and task length (due to viewing time variation) for
410 29 videos were coded, with an ICC(3,*k*) of 0.93 and 0.99 respectively, indicating excellent
411 reliability (Koo & Li, 2016).

412 *Models.* In four cases where sessions were repeated, the initial sessions were removed
413 from further analysis. First, we examined whether taking part in the touch screen test had an
414 effect on looking time in the free view task. As fewer monkeys had seen the images before (*n*
415 = 10) than those that had not (*n* = 19), we used a linear mixed model with log-transformed
416 looking time as the dependent variable and ID as a random effect. We made a binary variable
417 that split monkeys into those that had or had not seen the pair of images used for the free view
418 task previously, and examined it as a fixed effect in the model. Monkeys that had seen the
419 images before had significantly reduced looking times to images in the free view task

420 (experienced monkeys: $M = 4.43$ s ($SD = 3.41$); naïve monkeys: $M = 9.83$ s ($SD = 6.97$); $b = -$
421 0.95 , $SE = 0.27$, $p < .001$, 95% CI = $[-1.45, -0.42]$), indicating higher novelty of the images to
422 naïve subjects (Figure 3).

423 In a parallel linear mixed model, we examined fixed effects of image type and order of
424 presentation, as well as their interaction, on looking time. There was no difference in looking
425 time to younger and older images (young LT: $M = 8.28$ s ($SD = 6.56$); old LT: $M = 7.66$ s (SD
426 $= 6.52$); $b = -0.19$, $SE = 0.35$, 95% CI = $[-0.88, -0.48]$), but there was an effect of viewing
427 order, with looking time decreasing across sessions (first session: $M = 8.74$ s ($SD = 5.88$);
428 second session: $M = 7.20$ s ($SD = 7.06$); $b = -0.76$, $SE = 0.35$, 95% CI = $[-1.42, -0.11]$). The
429 interaction between image type and order viewed was not significant ($b = 0.71$, $SE = 0.61$,
430 95% CI = $[-0.46, 1.87]$).

431 Finally, we examined whether looking time differed between image type for naïve
432 monkeys only ($n = 19$), with a paired t-test. Despite differences in looking time between naïve
433 and experienced subjects, when we examined differences in looking time to each image for
434 naïve monkeys only, there was still no significant difference ($t(18) = -0.29$, $p = 0.77$). These
435 findings thus reflect findings for looking time from the touch screen studies.

436

437 ----- Table 3 here -----

438

439 **Discussion**

440 We examined whether looking time is an indication of preference, by measuring image choice
441 explicitly in relation to looking time responses. The results provide mixed evidence for a
442 relationship between looking time and preference. On the one hand, there was no relationship
443 between tasks that assessed (1) looking time to, and (2) choices between images of social
444 content. This lack of relationship is likely explained by the absence of variance in responses –

445 there were no patterns in either task indicating visual bias or preference for any of the social
446 categories. On the other hand, the training data revealed patterns of initial visual bias, and
447 overall preference, for images of objects, and especially food, over landscapes. Differences
448 for image choice between food and landscape held when accounting for adjustment in the
449 training procedures for the choice task. The results also revealed a choice preference for food
450 over object images, however this was not reflected in the looking time data. Since food and
451 landscape images represent the highest and lowest preferred categories respectively, this
452 suggests that links to visual bias may only be found for image categories with high preference
453 differences. Overall, these results suggest that visual bias to images of food over landscape,
454 and to a lesser extent images of objects over landscape, does reflect preferences when
455 presented with an explicit choice task.

456

457 ----- Figure 3 here -----

458

459 The fact that looking time in viewing task II did not decrease overall across the testing
460 sessions, and that responses by naïve subjects in their home environment elicited similar
461 responses as experienced monkeys, suggests that these findings are not a result of the
462 methodology used, such as presenting the images on a screen or in succession. The free view
463 task, however, did reveal order effects, in that monkeys looked longer at whichever image
464 they viewed first. We also found that monkeys looked at images longer when they were
465 novel, suggesting that previous experience influences gaze behavior. This latter finding is not
466 likely explained by differences in exposure to the researchers, since testing always took place
467 adjacent to the home enclosure, and observational data collection as well as target-training
468 were carried out in the home enclosure, giving all monkeys equal chance of exposure to the
469 researchers. Why the monkeys showed no category-level biases towards the social stimuli, in

470 contrast to earlier studies, (Almeling et al., 2016; Gothard et al., 2004; Pfefferle et al., 2014;
471 Schell et al., 2011) remains unclear, but may be due to differences found at the individual
472 rather than species level. Due to our small sample size in the touchscreen studies, we were
473 unable to explore whether this was the case.

474 Although our results tentatively suggest that visual bias in a viewing task could
475 indicate preference in an explicit choice task, it is necessary to understand why attending to
476 one stimulus more than another could be meaningful to the study individual or species. This is
477 important if gaze data is to be interpreted from an ecologically valid perspective (D'Eath,
478 1998; Morton et al., 2016). Indeed, as Aslin (2007) points out, one should also consider that a
479 lack of bias towards one stimulus-type does not mean that subjects do not discriminate
480 between stimuli. When using looking time as an experimental measure, we encourage the use
481 of clear linking hypotheses between visual bias and the underlying cognitive mechanisms.
482 This should include paying particular consideration to the social context that is being
483 assessed, since this is likely to drive gaze allocation.

484 To address the issue of ecological validity, one approach would be to examine visual
485 bias in parallel with other responses. For example, red-fronted lemurs spend more time
486 looking at and more time sniffing images of conspecifics compared with heterospecifics
487 (Rakotonirina et al., 2018). A number of studies have focused only on behavioral responses to
488 measure image discrimination. In goats (*Capra hircus*), ears are positioned forward for longer
489 in response to conspecific negative over positive facial expressions (Bellegarde et al., 2017),
490 which could be an adaptive response to potential threat (Bethell et al., 2012). Goats were also
491 found to direct first interactions towards positive over negative human facial expressions
492 (Nawroth et al., 2018). Discus fish (*Symphysopdon aequifasciatus*) perform partner specific
493 displays to lateral images of their mate, but not to images of unfamiliar conspecifics,
494 suggesting that they use facial color patterns to differentiate between conspecifics (Satoh

495 2016). Similarly, female jumping spiders (*Maevia inclemens*) were found to exhibit sexual
496 receptivity more often towards video of normal male morphs than digitally altered morphs
497 (Clark & Uetz, 1993). Approach behavior has additionally been used as a way to measure
498 response to facial stimuli in both brown capuchins (Morton et al., 2016) and horses (*Equus*
499 *caballus*; (Wathan, Proops, Grounds, & McComb, 2016). These measures may be more
500 informative to answer questions about stimuli recognition, avoidance or preference, and in
501 future research should be considered either alongside or instead of looking time paradigms.

502 An additional avenue of research that now shows promise for a number of species,
503 such as primates (Hopper et al., 2020) and dogs (Somppi, Törnqvist, Hänninen, Krause, &
504 Vainio, 2012), is the use of eye tracking. Eye tracking can be conducted non-invasively and
505 thus is of increasing application to a wider scope of research setups such as in zoos (Hopper et
506 al., 2020). The benefit of this approach is that it allows one to examine in detail gaze
507 responses to social stimuli, such as how species explore facial images (Kano, Call, &
508 Tomonaga, 2012; Wilson et al., 2020). This could be particular useful in instances such as
509 that presented in the current study, where we found no visual bias amongst categories of
510 different social stimuli. An eye tracking approach in this case could instead reveal differences
511 in tracking patterns between such stimuli, and could be more informative in highlighting
512 attentional differences at the individual level.

513 Regarding our own methodology, there are several limitations to address. First, in
514 viewing task I, the use of the reward randomization did not, as we hoped, dissociate the food
515 incentive from viewing the images. This did not detract however from initial attention to the
516 images. Second, we made several procedural adjustments during training for the choice task.
517 Sensitivity analyses however revealed that these changes had, at best, marginal effect on
518 overall choices. An additional limitation is that although our training data suggest a
519 relationship between looking time and choice, the use of different training stimuli meant we

520 were unable to test this directly (since these analyses were unplanned). Moreover, training on
521 the choice task took place after training on the initial viewing task, so we could not
522 counterbalance task order, although, the fact that different image sets were used in each task
523 means that this did not affect image novelty. Finally, in this study we focused on the
524 assessment of response to static images. Further investigation of looking time in response to
525 dynamic stimuli would be beneficial in understanding the relationship between gaze and
526 preference.

527 To summarize, the results presented in this study do suggest a possible link between
528 visual attention and preference. However, this study also revealed that common assumptions
529 such as preferences for images of infants over older subjects are not always borne out. We
530 therefore urge caution in the use of looking time measures as a way to assess social
531 preferences, particularly without validation from other measures. We encourage future
532 research to consider drawing more strongly on behavioral measures, or where possible eye
533 tracking, that could be more informative to the questions surrounding social perceptions and
534 preferences.

535

536 **Acknowledgements**

537 The authors acknowledge funding from the Ministry for Science and Education of Lower
538 Saxony and the Volkswagen Foundation through the program “Förderung exzellenter
539 Forschung in Niedersachsen: Forschungsverbund” Zuwendungsnummer VWZN3106, as well
540 as the Leibniz Association through funding for the Leibniz ScienceCampus Primate
541 Cognition. Thanks to L. Almeling and E. Rathke for providing images of Barbary macaques,
542 and to E. Rathke for providing equipment. Thanks also to H. Otto, D. Fuchs, B. Walter and L.
543 Ehrenreich for video coding and/or technical assistance, and members of the Cognitive
544 Ethology lab for discussions about methods and providing images of the monkeys, in

545 particular C. Schloegl and A. Tramacere. Thanks to R. Brockhausen for XBI technical
546 assistance, the keepers, N. Rabah and H. Mascher, for their support, and finally Michael
547 Beran and several anonymous reviewers for helpful comments on earlier versions of this
548 manuscript.

549

550 **Conflict of Interest**

551 The authors declare that they have no conflict of interest.

552

553 **Ethics**

554 This study was non-invasive and is in accordance with the German legal and ethical
555 requirements of appropriate animal procedures using nonhuman primates. As confirmed by
556 the Lower Saxony State Office for Consumer Protection and Food Safety, these experiments
557 do not constitute a procedure according to the animal welfare legislation (§7, Abs. 2
558 TierSchG); therefore a permit was not required (LAVES Document 33.19-42502-04).
559 Institutional approval was provided by the German Primate Center Animal Welfare Body
560 under application no. E2-18.

561

562

563 **References**

- 564
- 565 Adams, L. C., & Macdonald, S. E. (2018). Spontaneous preference for primate photographs in
566 Sumatran orangutans (*Pongo abelii*). *International Journal of Comparative Psychology*,
567 *31*, 1–16.
- 568 Albuquerque, N., Guo, K., Wilkinson, A., Savalli, C., Otta, E., & Mills, D. (2016). Dogs
569 recognize dog and human emotions. *Biology Letters*, *12*, 20150883.
570 <https://doi.org/10.1007/s10071-010-0316-y>
- 571 Almeling, L., Hammerschmidt, K., Sennhenn-Reulen, H., Freund, A. M., & Fischer, J.
572 (2016). Motivational shifts in aging monkeys and the origins of social selectivity.
573 *Current Biology*, *26*(13), 1744–1749. <https://doi.org/10.1016/j.cub.2016.04.066>
- 574 Aslin, R. N. (2007). What’s in a look? *Developmental Science*, *10*(1), 48–53.
- 575 Bellegarde, L. G. A., Haskell, M. J., Duvaux-ponter, C., Weiss, A., Boissy, A., & Erhard, H.
576 W. (2017). Face-based perception of emotions in dairy goats. *Applied Animal Behaviour*
577 *Science*, *193*, 51–59. <https://doi.org/10.1016/j.applanim.2017.03.014>
- 578 Bethell, E. J., Holmes, A., MacLarnon, A., & Semple, S. (2012). Evidence that emotion
579 mediates social attention in rhesus macaques. *PLoS ONE*, *7*(8).
580 <https://doi.org/10.1371/journal.pone.0044387>
- 581 Bird, C. D., & Emery, N. J. (2008). Using video playback to investigate the social preferences
582 of rooks, *Corvus frugilegus*. *Animal Behaviour*, *76*, 679–687.
583 <https://doi.org/10.1016/j.anbehav.2008.04.014>
- 584 Bolker, B. (2019). Linear Mixed-Effects Models using “Eigen” and S4.
- 585 Calapai, A., Berger, M., Niessing, M., Heisig, K., Brockhausen, R., Treue, S., & Gail, A.
586 (2016). A cage-based training, cognitive testing and enrichment system optimized for
587 rhesus macaques in neuroscience research. *Behavior Research Methods*, *5*, 1–11.
588 <https://doi.org/10.3758/s13428-016-0707-3>
- 589 Clark, D. L., & Uetz, G. W. (1993). Signal efficacy and the evolution of male dimorphism in
590 the jumping spider, *Maevia inclemens*. *Proceedings of the National Academy of Sciences*
591 *of the United States of America*, *90*, 11954–11957.
- 592 Costa, M., Gomez, A., Barat, E., Lio, G., Duhamel, J.-R., & Sirigu, A. (2018). Implicit
593 preference for human trustworthy faces in macaque monkeys. *Nature Communications*,
594 *9*, 4529. <https://doi.org/10.1038/s41467-018-06987-4>
- 595 D’Eath, R. B. (1998). Can video images imitate real stimuli in animal behaviour experiments?
596 *Biological Reviews*, *73*(3), 267–292.
- 597 Dubuc, C., Allen, W. L., Maestripieri, D., & Higham, J. P. (2014). Is male rhesus macaque
598 red color ornamentation attractive to females? *Behavioral Ecology and Sociobiology*,
599 *68*(7), 1215–1224. <https://doi.org/10.1007/s00265-014-1732-9>
- 600 Dufour, V., Pascalis, O., & Petit, O. (2006). Face processing limitation to own species in
601 primates, a response to social needs? *Behavioural Processes*, *73*, 107–113.
- 602 Fagot, J. (2000). *Picture perception in animals*. Psychology Press Ltd.
- 603 Fagot, J., Thompson, R. K. R., & Parron, C. (2010). How to read a picture: Lessons from
604 nonhuman primates. *PNAS*, *107*(2), 519–520. <https://doi.org/10.1073/pnas.0913577107>
- 605 Fantz, R. L. (1958). Pattern vision in young infants. *The Psychological Record*, *8*, 43–47.
- 606 Farroni, T., Menon, E., & Johnson, M. H. (2006). Factors influencing newborns’ preference
607 for faces with eye contact. *Journal of Experimental Child Psychology*, *95*, 298–308.
608 <https://doi.org/10.1016/j.jecp.2006.08.001>
- 609 Fujita, K. (1987). Species recognition by five macaque monkeys. *Primates*, *28*(3), 353–366.
- 610 Gothard, K. M., Erickson, C. A., & Amaral, D. G. (2004). How do rhesus monkeys (*Macaca*
611 *mulatta*) scan faces in a visual paired comparison task? *Animal Cognition*, *7*, 25–36.
612 <https://doi.org/10.1007/s10071-003-0179-6>

- 613 Hopper, L. M., Gulli, R. A., Howard, L. H., Kano, F., Krupenye, C., Ryan, A. M., & Paukner,
614 A. (2020). The application of noninvasive, restraint-free eye-tracking methods for use
615 with nonhuman primates. *Behavior Research Methods*, 1–28.
- 616 Kano, F., Call, J., & Tomonaga, M. (2012). Face and eye scanning in gorillas (*Gorilla*
617 *gorilla*), orangutans (*Pongo abelii*), and humans (*Homo sapiens*): Unique eye-viewing
618 patterns in humans among hominids. *Journal of Comparative Psychology*, 126(4), 388–
619 398. <https://doi.org/10.1037/a0029615>
- 620 Kim, Y., Choi, Y.-K., & Emery, S. (2013). Logistic regression with multiple random effects:
621 a simulation study of estimation methods and statistical packages. *Journal of the*
622 *American Statistical Association*, 67(3).
623 <https://doi.org/doi:10.1080/00031305.2013.817357>
- 624 Koba, R., & Izumi, A. (2008). Japanese monkeys (*Macaca fuscata*) discriminate between
625 pictures of conspecific males and females without specific training. *Behavioural*
626 *Processes*, 79, 70–73. <https://doi.org/10.1016/j.beproc.2008.04.005>
- 627 Kohda, M., Jordan, L. A., Hotta, T., Kosaka, N., Karino, K., Tanaka, H., ... Takeyama, T.
628 (2015). Facial recognition in a group-living cichlid fish. *PLoS ONE*, 10(11), e0142552.
629 <https://doi.org/10.1371/journal.pone.0142552>
- 630 Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation
631 coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163.
632 <https://doi.org/10.1016/j.jcm.2016.02.012>
- 633 Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate
634 that other individuals will act according to false beliefs. *Science*, 354(6308), 110–114.
635 <https://doi.org/10.1126/science.aaf8110>
- 636 Lyons, M., Marcinkowska, U., Moisey, V., & Harrison, N. (2016). The effects of resource
637 availability and relationship status on women’s preference for facial masculinity in men:
638 An eye-tracking study. *Personality and Individual Differences*, 95, 25–28.
639 <https://doi.org/10.1016/j.paid.2016.02.025>
- 640 Méary, D., Wu, L., Zhinan, L., Guo, K., & Pascalis, O. (2014). Seeing two faces together:
641 Preference formation in humans and rhesus macaques. *Animal Cognition*, 17, 1107–
642 1119. <https://doi.org/10.1007/s10071-014-0742-3>
- 643 Morton, F. B., Brosnan, S. F., Prétôt, L., Buchanan-Smith, H. M., O’Sullivan, E., Stocker, M.,
644 ... Wilson, V. A. D. (2016). Using photographs to study animal social cognition and
645 behaviour: Do capuchins’ responses to photos reflect reality? *Behavioural Processes*,
646 124, 38–46. <https://doi.org/10.1016/j.beproc.2015.10.005>
- 647 Nawroth, C., Albuquerque, N., Savalli, C., Single, M., & McElligott, A. G. (2018). Goats
648 prefer positive human emotional facial expressions. *Royal Society Open Science*,
649 5(180491).
- 650 Overduin-de Vries, A. M., Bakker, F. A. A., Spruijt, B. M., & Sterck, E. H. M. (2016). Male
651 long-tailed macaques (*Macaca fascicularis*) understand the target of facial threat.
652 *American Journal of Primatology*, 78(December 2014), 720–730.
653 <https://doi.org/10.1002/ajp.22536>
- 654 Paukner, A., Wooddell, L. J., Lefevre, C. E., Lonsdorf, E., & Lonsdorf, E. (2017). Do
655 capuchin monkeys (*Sapajus apella*) prefer symmetrical face shapes? *Journal of*
656 *Comparative Psychology*, 131(1), 73.
- 657 Pfefferle, D., Kazem, A. J. N., Brockhausen, R. R., Ruiz-lambides, A. V., & Widdig, A.
658 (2014). Monkeys spontaneously discriminate their unfamiliar paternal kin under natural
659 conditions using facial cues. *Current Biology*, 24, 1806–1810.
660 <https://doi.org/10.1016/j.cub.2014.06.058>
- 661 Rakotonirina, H., Kappeler, P. M., & Fichtel, C. (2018). The role of facial pattern variation
662 for species recognition in red-fronted lemurs (*Eulemur rufifrons*). *BMC Evolutionary*

- 663 *Biology*, 18, 1–10. <https://doi.org/10.1186/s12862-018-1126-0>
- 664 Rudolf Von Rohr, C., Van Schaik, C. P., Kissling, A., & Burkart, J. M. (2015). Chimpanzees’
665 bystander reactions to infanticide an evolutionary precursor of social norms? *Hum Nat*,
666 26, 143–160. <https://doi.org/10.1007/s12110-015-9228-5>
- 667 Sato, A., Koda, H., Lemasson, A., Nagumo, S., & Masataka, N. (2012). Visual recognition of
668 age class and preference for infantile features: Implications for species-specific vs
669 universal cognitive traits in primates. *PLoS ONE*, 7(5), e38387.
670 <https://doi.org/10.1371/journal.pone.0038387>
- 671 Satoh, S., Tanaka, H., & Kohda, M. (2016). Facial recognition in a Discus fish (Cichlidae):
672 experimental approach using digital models. *PLoS One*, 11(5), 1–11.
673 <https://doi.org/10.1371/journal.pone.0154543>
- 674 Schell, A., Rieck, K., Schell, K., Hammerschmidt, K., & Fischer, J. (2011). Adult but not
675 juvenile Barbary macaques spontaneously recognize group members from pictures.
676 *Animal Cognition*, 14, 503–509. <https://doi.org/10.1007/s10071-011-0383-8>
- 677 Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and
678 influences preference. *Nature Neuroscience*, 6(12), 1317–1322.
679 <https://doi.org/10.1038/nn1150>
- 680 Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability.
681 *Psychological Bulletin*, 86(2), 420–428.
- 682 Somppi, S., Törnqvist, H., Hänninen, L., Krause, C., & Vainio, O. (2012). Dogs do look at
683 images: eye tracking in canine cognition research. *Animal Cognition*, 15, 163–174.
684 <https://doi.org/10.1007/s10071-011-0442-1>
- 685 Waitt, C., & Little, A. C. (2006). Preferences for symmetry in conspecific facial shape among
686 *Macaca mulatta*. *International Journal of Primatology*, 27(1), 133–145.
687 <https://doi.org/10.1007/s10764-005-9015-y>
- 688 Waitt, C., Little, A. C., Wolfensohn, S., Honess, P., Brown, A. P., Buchanan-Smith, H. M., &
689 Perrett, D. I. (2003). Evidence from rhesus macaques suggests that male coloration plays
690 a role in female primate mate choice. *Proceedings of the Royal Society B: Biological*
691 *Sciences*, 270, S144–S146. <https://doi.org/10.1098/rsbl.2003.0065>
- 692 Wathan, J., Proops, L., Grounds, K., & McComb, K. (2016). Horses discriminate between
693 facial expressions of conspecifics. *Scientific Reports*, 6(38322), 1–11.
694 <https://doi.org/10.1038/srep38322>
- 695 Wilson, V. A. D., Kade, C., Moeller, S., Treue, S., Kagan, I., & Fischer, J. (2020). Macaque
696 gaze responses to the Primatar: A virtual macaque head for social cognition research.
697 *Frontiers in Psychology*, 11(July), 1–13. <https://doi.org/10.3389/fpsyg.2020.01645>
- 698 Winters, S., Dubuc, C., & Higham, J. P. (2015). Perspectives: the looking time experimental
699 paradigm in studies of animal visual perception and cognition. *Ethology*, 121, 625–640.
700 <https://doi.org/10.1111/eth.12378>
- 701 Zhang, H., Lu, N., Feng, C., Thurston, S. W., Xia, Y., & Tu, X. M. (2011). On fitting
702 generalized linear mixed-effects models for binary responses using different statistical
703 packages. *Statistics in Medicine*, 30(20), 2562–2572.
- 704

Table 1. Viewing time in viewing task I to three non-social image categories in first training session

	<i>b</i>	<i>SE</i>	<i>0.835% CI</i>	<i>99.165% CI</i>	<i>Adjusted p</i>
Food versus landscape					
<i>Intercept</i>	0.09	0.15	-0.39	0.51	1.00
<i>Food</i>	0.73	0.16	0.27	1.19	< .001
Object versus landscape					
<i>Intercept</i>	0.08	0.16	-0.51	0.53	1.00
<i>Objects</i>	0.73	0.15	0.22	1.23	< .001
Object versus food					
<i>Intercept</i>	0.77	0.14	0.12	1.33	< .001
<i>Objects</i>	-0.002	0.12	-0.61	0.56	1.00

Significant values highlighted in bold.

Table 2. Image categories chosen across choice task I training sessions

	<i>OR</i>	<i>SE</i>	<i>1.25% CI</i>	<i>98.75% CI</i>	<i>p</i>
<i>Full model</i>					
<i>Intercept</i>	4.39	0.06	3.84	4.98	< .001
<i>Food vs. Landscape</i>	1.62	0.06	1.44	1.88	< .001
<i>Objects vs. Landscape</i>	1.39	0.06	1.20	1.58	< .001
<i>Subset model</i>					
<i>Intercept</i>	7.12	0.05	6.35	7.99	< .001
<i>Object vs. Food</i>	0.86	0.05	0.75	0.96	< .01

Image choices for food and objects both compared with landscape (top) and object compared with food (bottom). Significant values highlighted in bold. Confidence intervals and p values adjusted for multiple comparisons between models.

Table 3. Looking time for viewing task II per social image category, and in relation to image choice

	<i>b</i>	<i>SE</i>	<i>2.5% CI</i>	<i>97.5% CI</i>	<i>p</i>
Age: young versus old					
<i>Intercept</i>	0.45	0.11	0.005	0.81	< .001
<i>Young</i>	0.008	0.10	-0.43	0.45	0.94
Group: outgroup versus ingroup					
<i>Intercept</i>	0.34	0.14	-0.31	0.90	0.01
<i>Outgroup</i>	0.15	0.14	-0.48	0.86	0.27
Group: conspecifics versus heterospecifics					
<i>Intercept</i>	0.56	0.14	-0.19	1.16	< .001
<i>Conspecifics</i>	-0.18	0.15	-0.90	0.61	0.25
Group: nonkin versus kin					
<i>Intercept</i>	-0.12	0.18	-1.18	0.63	0.48
<i>Nonkin</i>	0.33	0.16	-0.55	1.32	0.04
Instances images chosen					
<i>Intercept</i>	-0.03	0.19	-1.32	1.10	0.89
<i>Instances chosen</i>	0.14	0.05	-0.26	0.50	< .01

First four models indicate relationship between looking time and image type, per social image category. Last model (in grey) indicates relationship between looking time per image type and instances each image type chosen, across categories. CIs do not support the p-value in this model, suggesting the model is not stable.



Figure 1. Tasks and apparatus. (A), (B) View of the testing enclosure and XBI. (C) Viewing task I on the Elo touch screen and (D) on the XBI. (E) Choice task; the monkey selects the target, then chooses one of two images presented; (F) Viewing task II; the monkey touches the target for a reward and then is presented with an image. (G) Free view task: (left) viewing apparatus; (middle) set up: (a) experimenter, (b) test monkey, (c) image held in frame, (d) camcorder; (right) participating monkey.

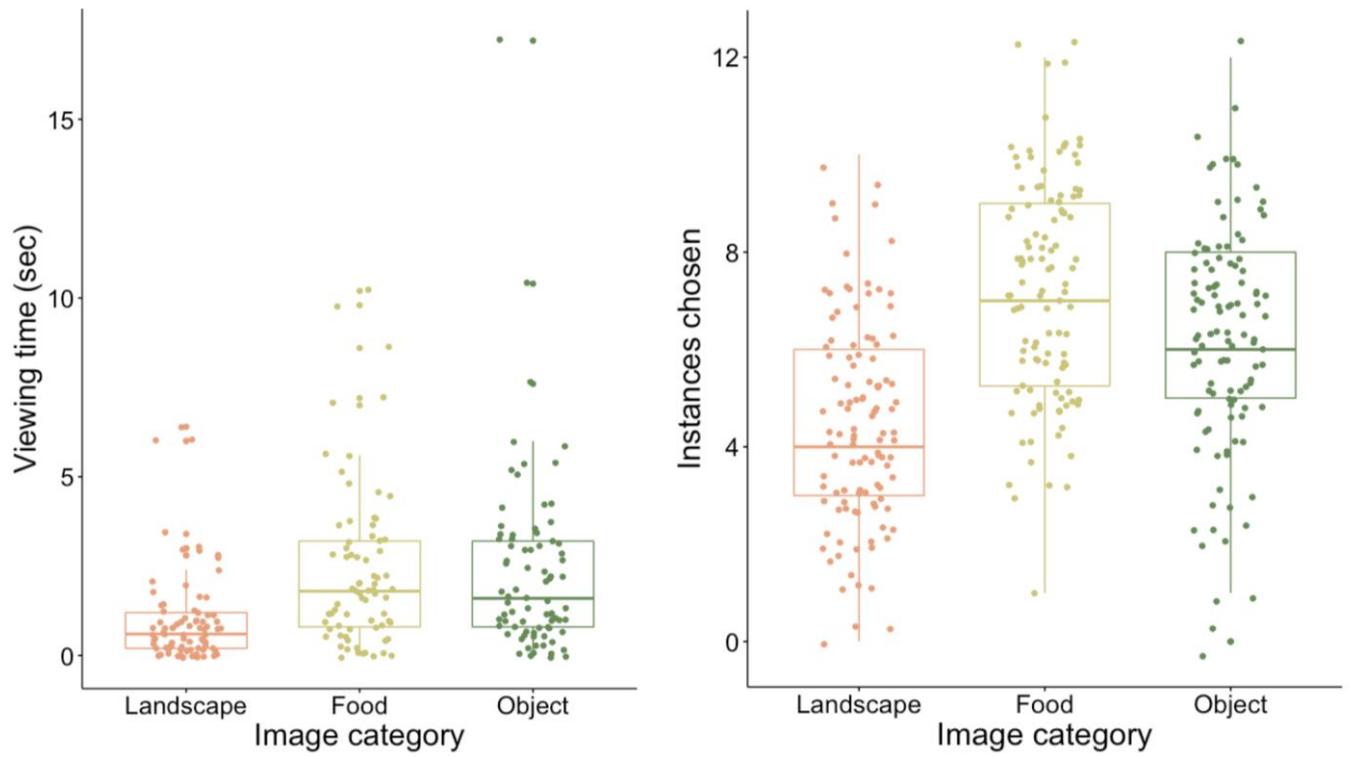


Figure 2. Response to non-social image categories. Left: looking time to images in first session of viewing task I. Right: images chosen across all training sessions of choice task I. Central horizontal bars indicate median values; upper and lower horizontal bars indicate the interquartile range (IQR); vertical error bars represent data points up to 1.5 x the IQR.

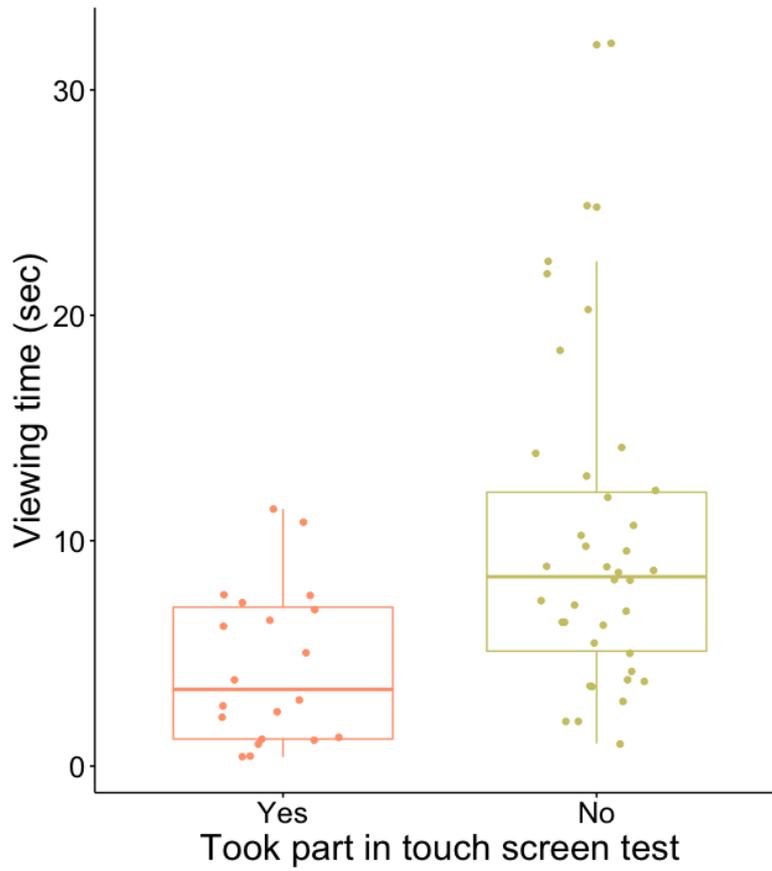


Figure 3. Looking time differences in the free view task between monkeys that did and did not see images in the touch screen test. Central horizontal bars indicate median values; upper and lower horizontal bars indicate the interquartile range (IQR); vertical error bars represent data points up to 1.5 x the IQR.

Appendix

Methods and Results

Experiment I.

Viewing task I.

Apparatus for training. For viewing task I, initial training for ten monkeys took place on an Elo touch screen, before being transferred to training on the XBI touch screen. As we used first training session to examine looking time to novel non-social images, data from eight of these monkeys was taken from sessions with the Elo touch screen (two were excluded as top positioned cameras were missing from their first session, and these were required for coding image type). This earlier training took place with the use of an Elo 17" SXGA TFT touchmonitor which was connected to an external MacBook Pro computer which ran on OS X El Capitan version 10.11.6 (see Figure 1). In this set up, cameras from the side and above filmed the monkeys during training, and monkeys were rewarded with cut raisins. The experimental software, MWorks, and the task, remained the same across both touch screen set-ups, with the exception that on the XBI, the task background was changed to white. Due to the lack of light in the XBI set-up, a white background enabled better vision of the monkey in the set-up.

Training stimuli. Landscape images were chosen from pre-existing footage taken by VW. Images were taken of the monkeys' food items, such as cut fruit and vegetables, and of novel household objects, including stationery, ornaments and kitchen utensils, which the monkeys were unlikely to have seen before. Novel objects were used to promote interest in the images.

Task limitations. Viewing task I was developed as a pilot method prior to the other tasks, to allow the monkeys to learn, through training with non-social images, to control how long they viewed each image for (similar to Fujita (1987) and Tanaka (2003)). We tried to test participants on this task. However, we were not successful, as the reward randomization did not work. In initial training sessions, monkeys took time to view the images and press the target. Over time, however, they learned to press the target repetitively and thus lost attention to the images, which meant that in later training sessions, it was not possible to code looking time to most of the images. Consequently, for this task we only present looking time from each monkey's initial training session when images were novel and they were clearly attending to them. This approach parallels viewing task II, where monkeys only saw each image once.

Choice task I.

Video coding. To assess monkeys' progress during training, all videos were coded by VW following each training session. This was done by playing videos in VLC and noting choices – including (non-social) image category chosen and side image chosen - in Microsoft Excel. We used this data to analyse the monkeys' choices across sessions. As the coder had to be familiar with the images in the training data in order to code the category correctly, obtaining a second coder for this was not possible.

Procedure adjustments. Due to their short attention span, this was a challenging task for the monkeys to learn. We thus adjusted some variables during training to make the task easier for the monkeys. The trial-start target was added after initial sessions (3 to 4 sessions each) with four monkeys, in order to separate the reward from the images; in initial sessions, reward was provided when the image pair appeared on screen, which distracted monkeys from choosing an image. The image pairs were initially presented for 20 s. If the monkeys did not select one of the two images within this time, a new trial would begin. However, this often led to the monkeys avoiding the stimuli and waiting for the next trial to begin automatically.

To prevent this, we increased the presentation time of the pair of pictures to a maximum of 300 s (this was implemented after an average of 5.88 sessions for eight of the participants and from the beginning for the other five participants). To encourage the monkeys to respond to the pairs presented, we also added a ‘red screen’ time-out for 3 s if they did not respond within the time limit (this was implemented after an average of 18.13 sessions for 8 monkeys and was not implemented for the other five monkeys). Once an image was selected, the enlarged image was initially set to remain for 10 s, but as the monkeys lost interest during this time, after three to four initial sessions in four subjects, we decreased it to 6 s.

Sensitivity analyses for procedure adjustments. Of the 2,220 trials analyzed, there were 200 trials when no start target was displayed; 540 trials when the image pair was displayed for 20 seconds instead of 300 seconds; 160 trials when the enlarged image was displayed for 10 instead of 6 seconds; and 420 trials when the red screen was used.

We ran two additional models to determine whether the procedural adjustments during training had any effect on the results. In the first model, we examined image choice in a subset of data across 7 monkeys using a Poisson model. This model focused on monkeys’ initial training sessions (mean number of sessions = 4.57 (SD = 5.13)), before any changes to procedure were implemented. Food and object images were chosen significantly above landscape images in these initial sessions (food: OR = 1.68, $SE = 0.11$, $p < .001$, 95% CI[1.34, 2.13]; object: OR = 1.59, $SE = 0.11$, $p < .001$, 95% CI[1.28, 2.04]; see Appendix Figure 1). The second model examined image choice across later trials, after changes were implemented (mean number of sessions = 8.29 (SD = 3.25)). Results showed a consistent pattern as with earlier trials (food: OR = 1.64, $SE = 0.08$, $p < .001$, 95% CI[1.40, 1.93]; objects: OR = 1.40, $SE = 0.08$, $p < .001$, 95% CI[1.21, 1.64]; see Appendix Figure 1). These findings are consistent with the full model, assessing response across all trials.

Experiment II.

Test stimuli. For the kin and age categories, as well as ingroup images, photos of familiar conspecifics were used. For outgroup conspecifics, we used images from a group of unfamiliar long-tailed macaques at the German Primate Center, to ensure all images were of equally unfamiliar individuals. For other species, we used images of Barbary macaques taken at the monkey park Forêt de Singes, Rocamadour, France.

When matching images into pairs, in some cases images were mirror-flipped to assume the same gaze direction as the matched pair ($n = 4$ for images of familiar conspecifics). The species and group categories contained only images of females with averted gaze. The age and kin categories contained both male and female images, with both direct and averted gaze. Where necessary, images were also adjusted for brightness, saturation and colour balance, to ensure minimal differences in lighting conditions when the images were taken. Images in the age condition were matched by individual, that is, we used images of current conspecifics, who were at either juvenile (1-3 years) or subadult (3-4.5 years) age, and matched these with images of the same individuals when they were infants (< 1 year). In the young group there was one exception; one individual featured was 23 months old. This image was matched with a subadult image, and the age difference between the image groups was not smaller than the age difference for the other images. All images in the older age group were taken an average of 2.03 years after the infant images, and had an average age of 2.83 years (SD = 0.90).

Viewing task II.

Sensitivity analyses. For viewing task II, in four cases, images of familiar conspecifics were mirror-flipped for testing to match the gaze direction of the paired image. To check whether this had any effect on looking time, we removed the cases of mirror-flipped images

and re-examined the data, which made no difference to looking time between image groups (Appendix Figure 2).

In the young age group, one individual was significantly older than the others photographed. We examined whether removing this image from the model had any effect on the difference in response to younger versus older. Removing images of the older individual from the young age category (eight cases) reduced the mean looking time for the young group from 1.68 s to 1.63 s, and had no significant effect on the model ($b = 0.07$, $SE = 0.20$, 95% CI = [-0.34, 0.52]). We also ran an additional analysis on a subset of the data for the age category, to account for differences between image pairs of difference between age groups. Three image pairs had a between-age group mean difference of 40.3 months. The other seven pairs had a between-age group mean difference of 18.6 months. We re-ran the generalized linear mixed model for the age category from the main analysis, examining looking time just for this subset of images. For the age category differences, initial examination of the mean looking times revealed that, for the seven image pairs with a smaller age gap between groups, mean looking time to the young group ($M = 1.72$ s) was almost the same as for the older age group ($M = 1.74$ s). For the three image pairs with a larger age gap between groups, mean looking time was higher for the young group ($M = 1.68$ s) than the old group ($M = 1.61$ s). The model revealed no significant difference in looking time between older and younger images, for the three image pairs with the larger age gap ($b = 0.05$, $SE = 0.21$, 95% CI = [-0.71, 0.78]; see Appendix Figure 3).

Finally, we additionally examined, using linear mixed models, whether differences in sex of photographed monkeys, or gaze direction of images, affected looking time, particularly as direct gaze may be perceived as threatening (de Waal, 1976). Looking time to the images did not differ with gaze direction of image, and was in the opposite direction from what one would expect (LT averted gaze: $M = 1.62$ s; LT direct gaze: $M = 1.71$ s; $b = 0.12$, $SE = 0.09$, 95% CI = [-0.06, 0.31]). Nor did looking time differ with sex of the monkey featured (LT female: $M = 1.64$ s; LT male: $M = 1.74$ s; $b = 0.03$, $SE = 0.14$, 95% CI = [-0.25, 0.33]).

Choice task II.

Sensitivity analyses. In addition to examining the relationship between image choice and looking time, we examined whether first touch – instances where monkeys lightly touched the image, but did not touch it enough to activate it to change – had a relationship with looking time. We ran a generalized linear mixed model for each image category with first image touched as the dependent binomial variable, log transformed viewing time and side image chosen as fixed effects and ID as a random effect. Consistent with results for which image chosen in each trial, first touch had no relationship to looking time for any of the categories (group: OR = 1.53, $SE = 0.51$, 95% CI[0.57, 4.15]; age: OR = 0.82, $SE = 0.30$, 95% CI[0.46, 1.48]; kin: OR = 0.56, $SE = 1.07$, 95% CI[0.07, 4.58]; species: OR = 1.15, $SE = 0.65$, 95% CI[0.32, 4.12]; see Appendix Table 2).

References

- Fujita, K. (1987). Species recognition by five macaque monkeys. *Primates*, 28(3), 353–366.
- Tanaka, M. (2003). Visual preference by chimpanzees (*Pan troglodytes*) for photos of primates measured by a free choice-order task: implication for influence of social experience. *Primates*, 44, 157–165. <https://doi.org/10.1007/s10329-002-0022-8>

Appendix Table 1. Demographics and participation of study participants across tasks

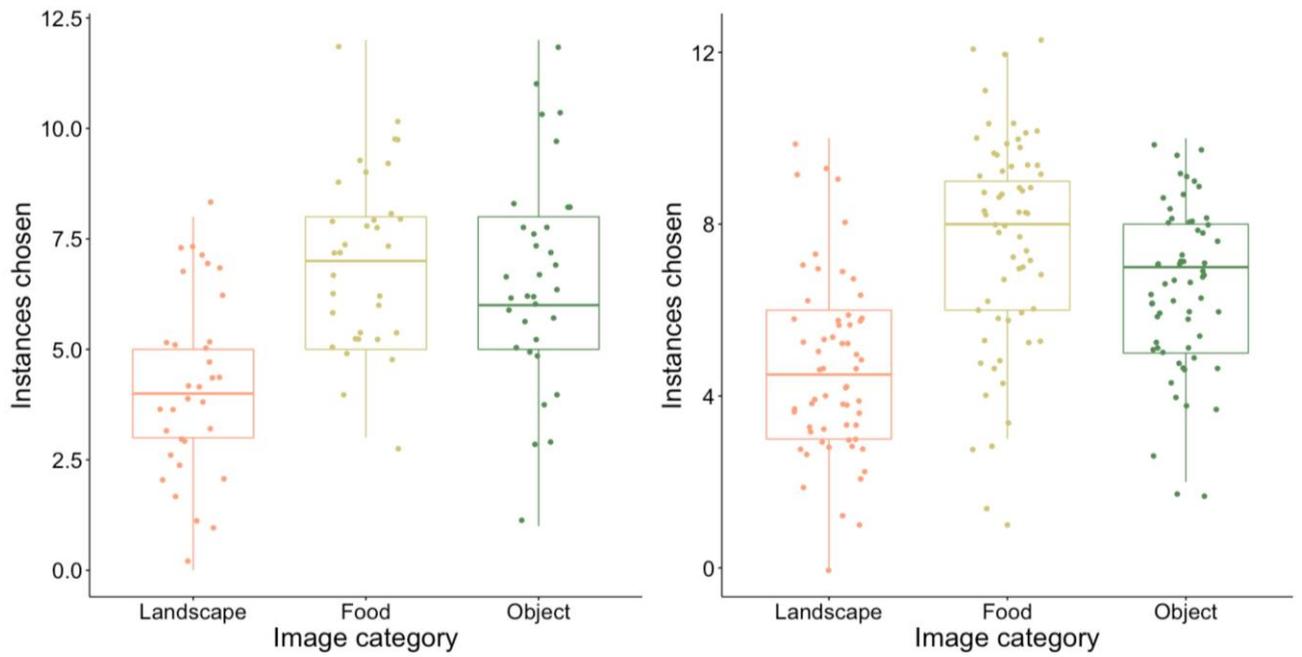
ID	Age at start of testing (yrs) ¹	Sex	VTI ELO training	VTI XBI		Choice		VTII	Free view
			Training	Training	Test	Training	Test	Test	Test
Natascha	31	F							Y
Indira	19	F							Y
Luka	18	F							Y
Sunny	15	M	1			17	Y	Y	Y
Inga	14	F							Y
Maja	10	F	11	2	Y	29			
Sally	10	F							Y
Sophie	8	F	3						Y
Lenny	9	M							Y
Leisa	8	F							Y
Ilja	5	M		6	Y	32	Y	Y	Y
Mila	5	F		6	Y	17	Y	Y	Y
Linus	4	M	6	6	Y	13	Y	Y	Y
Max	4	M	16			6	Y	Y	Y
Snickers	3	F	6	2	Y	11		Y	Y
Iowa	4	F							Y
Mars	3	M	16[†]	5	F	16	Y	Y	Y
Milka	3	F	24[†]			22			Y
Sissi	4	F							Y
Lukas	2	M		9				F	Y
Mara	2	F	7	7	Y*	43		Y	Y
Meiwi	2	F		14		42			Y
Sambia	2	F		9		42		Y	Y
Sverre	3	F							Y
Smilla	2	M							Y
Moritz	1	M	6	8		30	Y	Y	Y
Sherry	2	M							Y
Simon	1	F							Y
Madita	1	F							Y
Mona	1	F							Y

Note. Numbers depict no. training sessions prior to testing. VT = viewing task. Bold indicates counterbalance between viewing task II and choice task, i.e. which task monkey completed first. Y = completed test, F = failed to start or complete test. Bold-italics = viewed adult condition first. [†]Excluded from training data analysis, due to lack of downward angled camera in their first session, from which footage the coder could extract image type. *Tested in error. ¹Age calculated to the nearest year. Age determined from when touchscreen testing began. For monkeys that took part only in the free view task, age is calculated from when the free view task began.

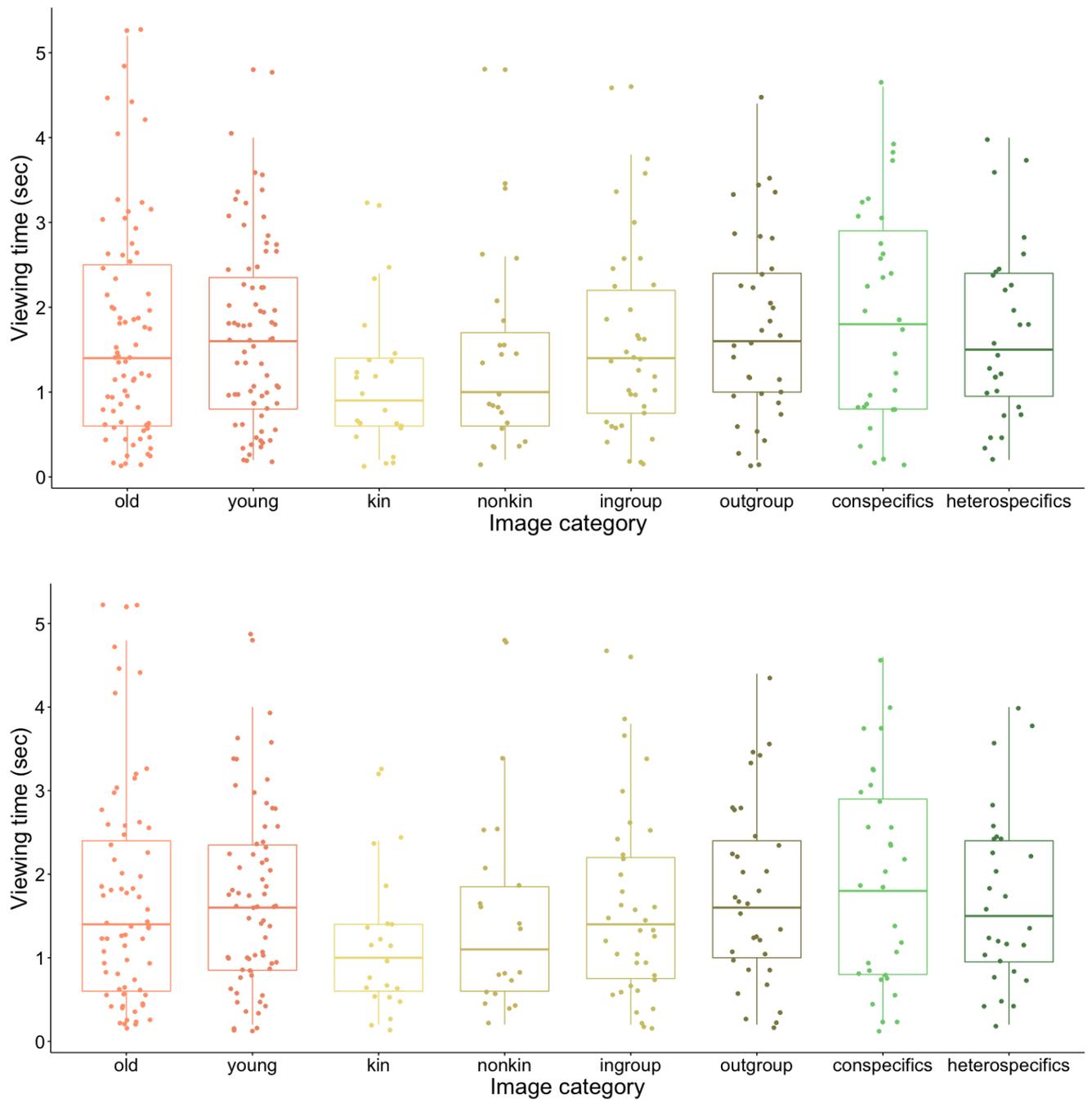
Appendix Table 2. Total number of instances of image choice and first touch from choice task II

		Image chosen					
Old	Young	Ingroup	Outgroup	Conspecifics	Heterospecifics	Kin	Nonkin
32	34	13	15	13	15	11	12
		First touch					
Old	Young	Ingroup	Outgroup	Conspecifics	Heterospecifics	Kin	Nonkin
25	25	10	15	12	11	10	8

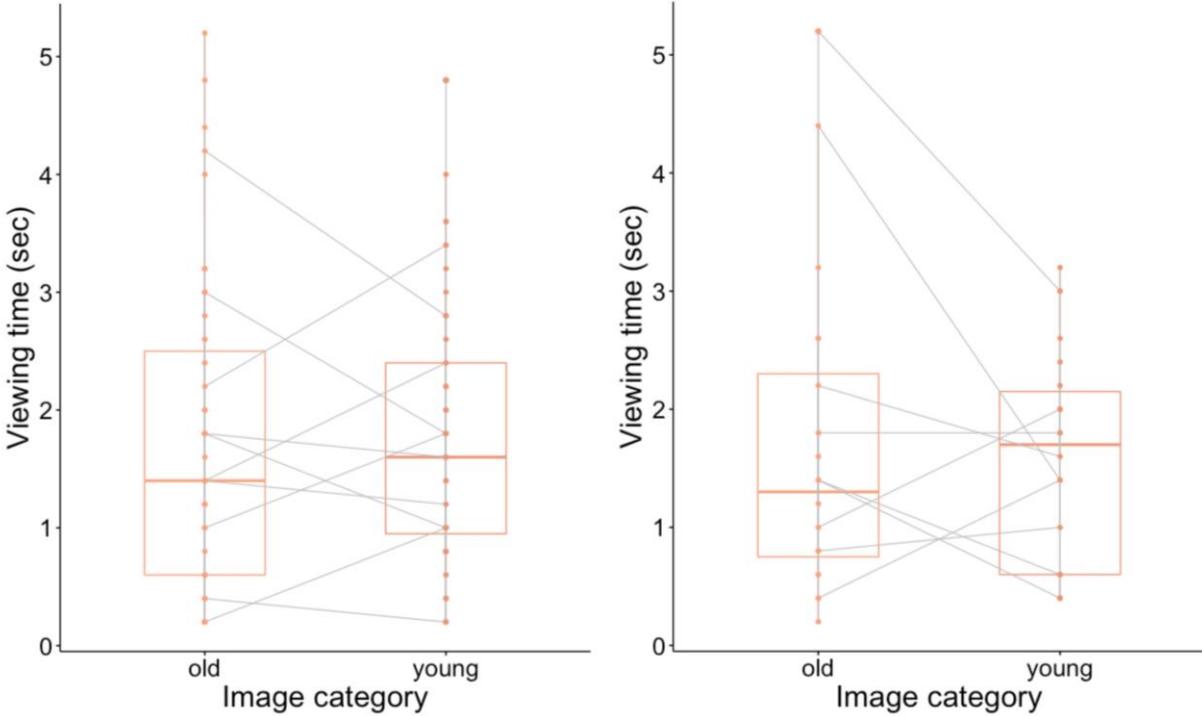
Note. Number of instances images chosen from each category (top) and number of instances when monkeys first touched an image but did not select it (i.e. image size did not increase).



Appendix Figure 1. Image preferences on choice task I, before (left) and after (right) procedural changes. Central horizontal bars indicate median values; upper and lower horizontal bars indicate the interquartile range (IQR); vertical error bars represent data points up to 1.5 x the IQR.



Appendix Figure 2. Looking time to social image categories on viewing task II, for all images (top) and with mirror-flipped images removed (bottom). Central horizontal bars indicate median values; upper and lower horizontal bars indicate the interquartile range (IQR); vertical error bars represent data points up to 1.5 x the IQR.



Appendix Figure 3. Looking time on viewing task II to age-category image pairs with different age gaps. Left: smaller age gap between younger and older groups, for seven image pairs. Right: larger age gap between younger and older groups, for three image pairs. Central horizontal bars indicate median values; upper and lower horizontal bars indicate the interquartile range (IQR); vertical error bars represent minimum and maximum values.