



Testing the Relationship Between Looking Time and Choice Preference in Long-tailed Macaques

Vanessa A. D. Wilson^{1,2,3,*}, Carolin Kade¹, and Julia Fischer^{1,2,3}

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 (Email: vanessa.wilson@unine.ch)

Citation – Wilson, V. A. D., Kade, C., & Fischer, J. (2021). Testing the relationship between looking time and choice preference in long-tailed macaques. *Animal Behavior and Cognition*, 8(3), 351-375. <https://doi.org/10.26451/abc.08.03.03.2021>

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

Keywords – *Macaca fascicularis*, Visual bias, Social cognition, Attention, Gaze

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

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

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

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

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

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

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

We collected data in three experimental phases, on five different tasks. Experiment 1 consisted of training monkeys on two different touchscreen tasks (viewing task 1 and choice task 1) using non-social images. Experiment 2 consisted of testing monkeys' responses to social images on two different touchscreen tasks (viewing task 2 and choice task 2). Experiment 3 consisted of collecting data on the free view task, where monkeys viewed images presented in their home enclosure. In all three viewing tasks, images were presented individually, and we measured looking time to each image, defined by the monkeys clearly orienting their head and eyes towards the images. The choice tasks presented the monkeys with pairs of images. In choice task 1, they had to learn to select one image to progress to the next trial and subsequently, the reward. We then tested their social preferences on this paradigm in choice task 2. For this we measured type, and side, of image chosen. All tasks except the free view task took place in test cubicles where monkeys interacted with a touch screen. The free view task took place in the monkeys' enclosure (see Table A1). Details of each task are provided below.

General Method

Participants

We collected data from a group of 36 long-tailed macaques (*Macaca fascicularis*) housed at the German Primate Center, Göttingen. The group had access to both an indoor (49 m²) and outdoor (141 m²) enclosure, with ad libitum access to food, water, and enrichment. Details of participating monkeys can be found in Table A1.

Testing Apparatus

We used the XBI touch screen – eXperimental Behavioural Instrument (Calapai et al., 2016), with a 15 in. Elo touchmonitor, MacBook Air (macOS Sierra Version 10.12.6), HD Wi-Fi camera with wide angle lens, and two pumps and drinking tube for automated fluid reward (see Figure 1). Experiments were run using MWorks (version 0.7). Individuals participated voluntarily in all touch screen testing, which took place in a separated testing area divided into six cubicles (height: 2.6 m; width: 2.25 m; depth: 1.25 m). Monkeys were rewarded for participation during testing with diluted grape juice (diluted at 200 ml juice:100 ml water; 0.25 ml per reward). Testing hours were Monday to Friday, 10:00 – 12:00 and 14:00 – 18:00.

Figure 1

Tasks and Apparatus



Note. Panels A, B: View of the testing enclosure and XBI. Panel C: Viewing task 1 on the Elo touch screen and Panel D: on the XBI. Panel E: Choice task; the monkey selects the target, then chooses one of two images presented; Panel F: Viewing task 2; the monkey touches the target for a reward and then is presented with an image. Panel 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.

Touchscreen Training

For the touch screen tasks, monkeys were trained and tested individually. All monkeys received prior training on the touch screen before commencing any task. When they exhibited excess caution towards the screen, for example by retreating when a stimulus appeared on-screen, we habituated them by pairing them with a partner who was comfortable interacting with the touch screen display.

Video Coding

Looking time (in viewing task 1, viewing task 2, and the free view task) as well as which side of the screen the monkey touched first, and which image they touched first (choice tasks 1 and 2) were coded from videos (image chosen was recorded by MWorks, but sometimes the monkeys' first touch to an image did not activate it, so we checked for this when coding the videos). We used the free behavioral coding software SolomonCoder (version 17.03.22). For each task (excluding choice task 1; see Appendix for details), a subset of videos was coded by two coders to assess interrater reliability. At least one of the coders was always naïve to the hypotheses.

General Analytical Procedure

Analyses were conducted in R studio version 1.0.153. We ran a series of mixed effects models using the lme4 package. For all models, we ran bootstrapped 95% confidence intervals. Where we adjusted models with a Bonferroni correction, we report the corresponding adjusted confidence intervals. For the viewing tasks, our analyses excluded images the monkeys did not look at. For viewing task 1, choice tasks 1 and 2, and visual bias to image category in viewing task 2, we set the number of adaptive Gauss-Hermite quadrature points to zero to aid model convergence. The number of Gauss-Hermite quadrature points (nAGQ) determines how the random effects are distributed in the model. When fitting glmms using the lme4 package in R, nAGQ defaults to one (Bolker, 2019; Zhang et al., 2011). Setting nAGQ to zero means that parameter estimation is less exact, but faster (Bolker, 2019), since the computational time decreases with fewer quadrature points (Kim et al., 2013). Each model is described in detail below.

Experiment 1

Viewing Task 1

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

Participants

We report data from 13 monkeys trained on this task (6M; 7F; M age = 5.91 years, SD = 5.47), eight of which received their initial training on a different apparatus (see Table A1 and Appendix for details).

Stimuli

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

Design and Procedure

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

Interrater Reliability

We calculated interrater reliability for looking time to each image for seven monkeys, using intraclass correlation coefficients (Koo & Li, 2016; Shrout & Fleiss, 1979). We estimated the reliability of mean values across k coders, ICC(3, k). For viewing task 1 we report an ICC(3, k) of .72, indicating good reliability (Koo & Li, 2016).

Models

To examine differences in initial looking time to each non-social image category, we ran three pair-wise generalized linear mixed models, fitted with a Gamma distribution, adjusted with Bonferroni correction. The models contrasted 1) food versus landscape, 2) landscape versus object, and 3) food versus object. Looking time was the dependent variable, image category was the binomial fixed effect and ID was a random effect.

Results

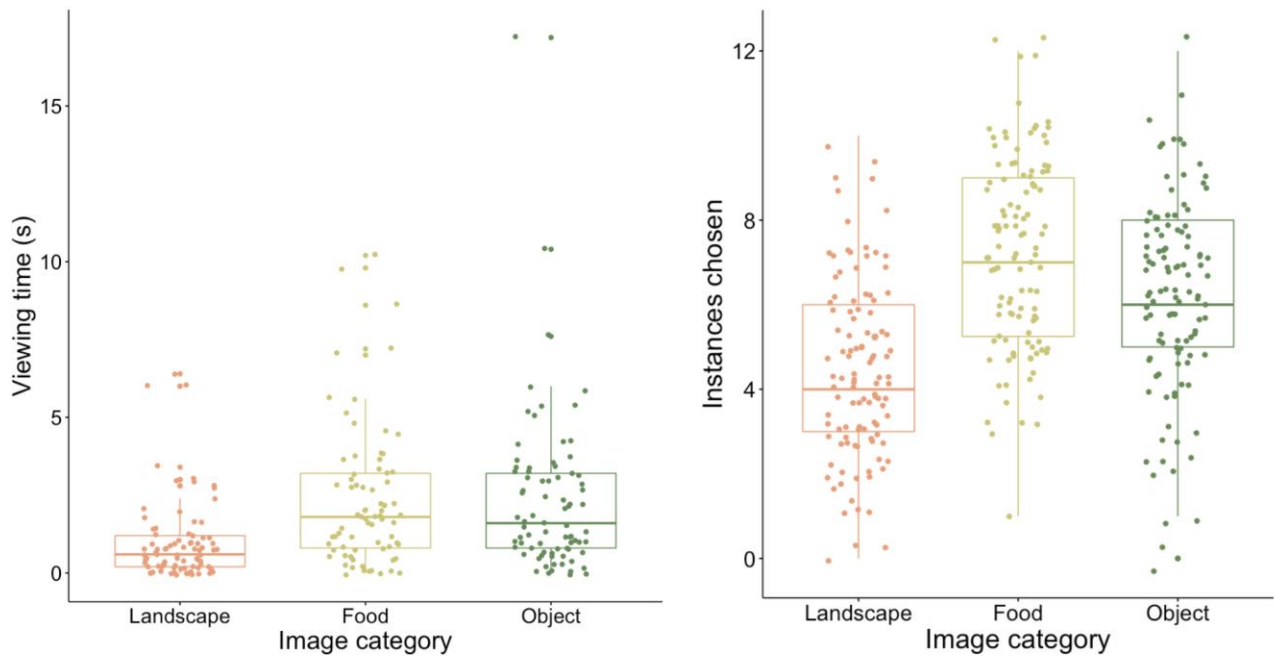
Looking time to images of food and objects was significantly longer than for landscapes, even after adjusting for multiple comparisons. There was no significant difference in looking time between objects and food (see Table 1 and Figure 2, left).

Table 1

Viewing Time in Viewing Task 1 to Three Non-Social Image Categories in the First Training Session

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

Note. Significant values highlighted in bold.

Figure 2*Response to Non-Social Image Categories*

Note. Left: looking time to images in first session of viewing task 1. Right: images chosen across all training sessions of choice task 1. 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.

Choice Task 1

Participants

Thirteen monkeys were trained on this task, eleven of which took part in viewing task 1 (6M, 7F; M age = 5.79 years, SD = 5.58; number training sessions: M = 18.71, SD = 9.23, range = 6–32 sessions).

Stimuli

Stimuli are as described for viewing task 1, except we used separate training sets for each task to ensure that images were novel and therefore promoted interest in the task.

Design and Procedure

We presented monkeys with pairs of images from which they could make a choice. Each trial started with a central target (green circle, 3.93 cm x 3.93 cm) on the screen, which provided reward when touched. Next, two images were presented on either side of the screen (7.85 cm x 7.85 cm). Order and side of image presentation were randomized. Participants had to select one image, though they were not rewarded for this – by providing juice only at the start of each trial (upon touching the target circle), reward was decoupled from image viewing. Image pairs were presented for up to 300 s, after which a new trial would start. When an image was selected, the other image would disappear and the selected image would increase in size (15.7 cm x 15.7 cm; see Figure 1) and remain for six seconds, before the next trial started.

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

Results

We examined image choice across training sessions for seven monkeys that passed training, using a generalized linear mixed Poisson model. We excluded any incomplete training sessions ($n = 21$ for five monkeys). Number of instances of each category chosen per monkey/session was the dependent variable, image category was the fixed effect and ID was the random effect. In a full model, we compared food and object choices over landscape. Consistent with looking time, monkeys chose food and object images more than landscape. A subset model that compared only food and object choices revealed that monkeys chose food images significantly more than object images (see Table 2 and Figure 2, right). Due to some procedural adjustments during training (see Appendix for details), we ran additional models to determine whether these changes had any overall effect on choice preferences. Results were consistent with our full model, reflecting preferences for food and object over landscape in both early and later trials (see Figure A1 and Appendix for details).

Table 2

Image Categories Chosen Across Choice Task 1 Training Sessions

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

Note. 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.

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

Discussion

As the images used for viewing task 1 and choice task 1 were, for the purposes of training, different, we could not directly examine the relationship between looking time and choice for these two tasks. The results do, however, indicate that preferences for images of food and objects in choice task 1 were consistent with longer looking times to novel food and object images in viewing task 1. This suggests that there may be a relationship between looking time and category preference for non-social images.

Experiment 2

Viewing task 2

Viewing task 2 was developed as an alternative testing procedure to that developed with viewing task 1. Although six monkeys reached criterion on viewing task 1, testing using this method was not feasible (see Appendix for details). Using data from viewing task 2, we examined the relationship between looking time to social images, and selection of social images in Choice task 2. Choice task 2 and viewing task 2 were counterbalanced in order of presentation.

Participants

Eleven monkeys were tested, but one monkey did not continue to engage the apparatus and thus failed to complete the test, so data were analyzed from ten monkeys, seven of which took part in viewing task 1 (6M, 4F; M age = 5.07, SD = 3.96).

Test Stimuli

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

Design and Procedure

The monkeys started each trial by pressing a target (purple diamond; 3.93 cm x 3.93 cm) to receive a reward. They were then presented with one image (15.7 cm x 15.7 cm; Figure 1) for six seconds, randomized in order. As the monkeys were familiar with the target-reward system, no training was required. Each monkey received one test session consisting of 36 – 46 trials, which varied due to number of kin images available per subject (between 18 and 23 image pairs: M = 20.71 trials, SD = 1.89).

Interrater Reliability

Viewing times were coded for interrater reliability from five videos. ICC(3, k), was .92, indicating excellent reliability (Koo & Li, 2016).

Results

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

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

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

Table 3*Looking Time for Viewing Task 2 Per Social Image Category, and in Relation to Image Choice*

	<i>b</i>	<i>SE</i>	2.5% CI	97.5% CI	<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	.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	.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	.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	.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

Note. 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.

Choice task 2

Participants

Seven monkeys that reached criterion in choice task 1 were tested on this task (5M; 2F; *M* age = 5.89 years, *SD* = 4.46). All seven monkeys also took part in viewing task 2.

Test Stimuli

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

Design and Procedure

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

Interrater Reliability

From six videos, reliability of side chosen and first touch were assessed by checking for discrepancies between coders. Three discrepancies for side chosen were found, only one of which

indicated an error by the primary coder. This was corrected in data analysis. Two discrepancies were found for first touch, neither of which were errors by the primary coder.

Results

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

To examine whether looking time in viewing task 2 predicted image choice in the choice task, we aggregated looking time to each image group per monkey, and paired this with the number of instances images in each group were chosen. We ran a generalized linear mixed model for each image category, with a Gamma distribution, looking time as the dependent variable, instances images were chosen as fixed effects and ID as a random effect. Image choice did not predict looking time to images in viewing task 2 across image categories (see Table 3, grey rows).

To examine whether there was any effect of test order (i.e., the order in which monkeys received the choice task 2 or viewing task 2) on response in the viewing task, we ran a generalized linear mixed model with a Gamma distribution, with looking time as the dependent variable, order as a fixed effect and ID as a random effect. Seeing the images in the choice task first did not have any effect on looking time to the images when receiving the viewing task second ($b = 0.43$, $SE = 0.28$, 95% CI [-0.47, 1.24]).

Experiment 3

Free View Task

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

Participants

We tested 29 monkeys (10M, 19F; M age = 7.22 years, $SD = 6.82$), 14 of which had participated in at least one of the touchscreen tasks, and ten of which had participated in viewing task 2.

Test Stimuli and Apparatus

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

Design and Procedure

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

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

Interrater Reliability

Viewing time and task length (due to viewing time variation) for 29 videos were coded, with an ICC(3,*k*) of .93 and .99 respectively, indicating excellent reliability (Koo & Li, 2016).

Results

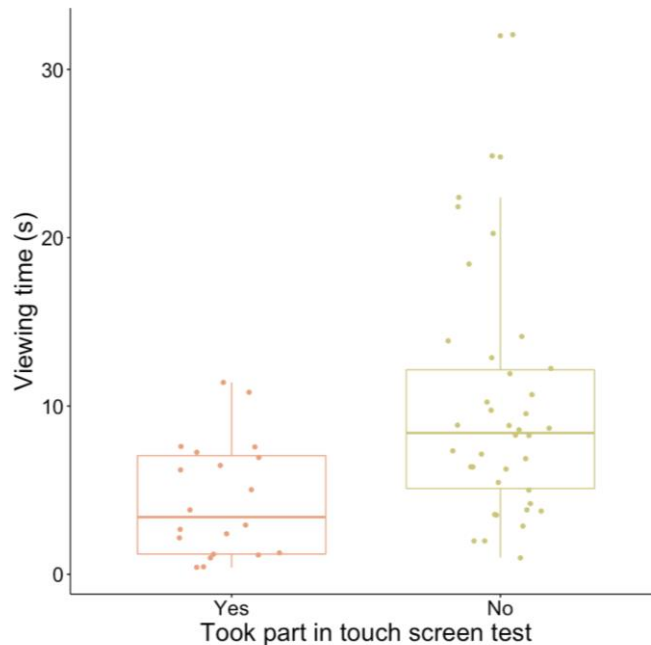
In four cases where sessions were repeated, the initial sessions were removed from further analysis. First, we examined whether taking part in the touch screen test had an effect on looking time in the free view task. As fewer monkeys had seen the images before ($N = 10$) than those that had not ($N = 19$), we used a linear mixed model with log-transformed looking time as the dependent variable and ID as a random effect. We made a binary variable that split monkeys into those that had or had not seen the pair of images used for the free view task previously and examined it as a fixed effect in the model. Monkeys that had seen the images before had significantly reduced looking times to images in the free view task (experienced monkeys: $M = 4.43$ s ($SD = 3.41$); naïve monkeys: $M = 9.83$ s ($SD = 6.97$); $b = -0.95$, $SE = 0.27$, $p < .001$, 95% CI [-1.45, -0.42]), indicating higher novelty of the images to naïve subjects (Figure 3).

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

Finally, we examined whether looking time differed between image type for naïve monkeys only (those that had not taken part in viewing task 2; $N = 19$), with a paired t-test. Despite differences in looking time between naïve and experienced subjects, when we examined differences in looking time to each image for naïve monkeys only, there was still no significant difference ($t(18) = -0.29$, $p = .77$). These findings for looking time match those from the touch screen studies.

Figure 3

Looking Time Differences in the Free View Task Between Monkeys That Did and Did Not See Images in the Touch Screen Test.



Note. 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.

General Discussion

We examined whether looking time is an indication of preference, by measuring image choice explicitly in relation to looking time responses. The results provided mixed evidence for a relationship between looking time and preference. On the one hand, there was no relationship between tasks that assessed (1) looking time to, and (2) choices between images of social content. This lack of relationship is likely explained by the absence of variance in responses – there were no patterns in either task indicating visual bias or preference for any of the social categories. On the other hand, the training data revealed patterns of initial visual bias, and overall preference, for images of objects, and especially food, over landscapes. Differences for image choice between food and landscape held when accounting for adjustment in the training procedures for the choice task. The results also revealed a choice preference for food over object images; however, this was not reflected in the looking time data. Since food and landscape images represent the highest and lowest preferred categories respectively, this suggests that links to visual bias may only be found for image categories with high preference differences. Overall, these results suggest that visual bias to images of food over landscape, and to a lesser extent images of objects over landscape, does reflect preferences when presented with an explicit choice task.

The fact that looking time in viewing task 2 did not decrease overall across the testing sessions, and that responses by naïve subjects in their home environment elicited similar responses as experienced monkeys, suggests that these findings are not a result of the methodology used, such as presenting the images on a screen or in succession. The free view task, however, did reveal order effects, in that monkeys looked longer at whichever image they viewed first. We also found that monkeys looked at images longer when they were novel, suggesting that previous experience influences gaze behavior. This latter finding is not likely explained by differences in exposure to the researchers, since testing always took place adjacent to the home enclosure, and observational data collection as well as target-training were carried out in the home enclosure, giving all monkeys equal chance of exposure to the researchers.

Why the monkeys showed no category-level biases towards the social stimuli, in contrast to earlier studies (Almeling et al., 2016; Gothard et al., 2004; Pfefferle et al., 2014; Schell et al., 2011) remains unclear, but may be due to differences found at the individual rather than species level. Due to our small sample size in the touchscreen studies, we were unable to explore whether this was the case.

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

To address the issue of ecological validity, one approach would be to examine visual bias in parallel with other responses. For example, red-fronted lemurs spend more time looking at and more time sniffing images of conspecifics compared with heterospecifics (Rakotonirina et al., 2018). A number of studies have focused only on behavioral responses to measure image discrimination. In goats (*Capra hircus*), ears are positioned forward for longer in response to conspecific negative over positive facial expressions (Bellegarde et al., 2017), which could be an adaptive response to potential threat (Bethell et al., 2012). Goats were also found to direct first interactions towards positive over negative human facial expressions (Nawroth et al., 2018). Discus fish (*Symphysopdon aequifasciatus*) perform partner specific displays to lateral images of their mate, but not to images of unfamiliar conspecifics, suggesting that they use facial color patterns to differentiate between conspecifics (Sato et al., 2016). Similarly, female jumping spiders (*Maevia inclemens*) were found to exhibit sexual receptivity more often towards video of normal male morphs than digitally altered morphs (Clark & Uetz, 1993). Approach behavior has been used as a way to measure response to facial stimuli in both brown capuchins (Morton et al., 2016) and horses (*Equus caballus*; Wathan et al., 2016). These measures may be more informative to answer questions about stimuli recognition, avoidance, or preference, and in future research should be considered either alongside or instead of looking time paradigms.

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

Regarding our own methodology, there are several limitations to address. First, in viewing task 1, the use of the reward randomization did not, as we hoped, dissociate the food incentive from viewing the images. This did not detract however from initial attention to the images. Second, we made several procedural adjustments during training for the choice task. Sensitivity analyses however revealed that these changes had, at best, marginal effect on overall choices. An additional limitation is that although our training data suggest a relationship between looking time and choice, the use of different training stimuli meant we were unable to test this directly (since these analyses were unplanned). Moreover, training on the choice task took place after training on the initial viewing task, so we could not counterbalance task order, although, the fact that different image sets were used in each task means that this did not affect image novelty. Finally, in this study we focused on the assessment of response to static images. Further investigation of looking time in response to dynamic stimuli would be beneficial in understanding the relationship between gaze and preference.

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

Acknowledgements

The authors acknowledge funding from the Ministry for Science and Education of Lower Saxony and the Volkswagen Foundation through the program “Förderung exzellenter Forschung in Niedersachsen: Forschungsverbund” Zuwendungsnummer VWZN3106, as well as the Leibniz Association through funding for the Leibniz-ScienceCampus Primate Cognition. Thanks to L. Almeling and E. Rathke for providing images of Barbary macaques, and to E. Rathke for providing equipment. Thanks also to H. Otto, D. Fuchs, B. Walter and L. Ehrenreich for video coding and/or technical assistance, and members of the Cognitive Ethology lab for discussions about methods and providing images of the monkeys, in particular C. Schloegl and A. Tramacere. Thanks to R. Brockhausen for XBI technical assistance, the keepers, N. Rabah and H. Mascher, for their support, and finally several anonymous reviewers for helpful comments on earlier versions of this manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest.

Ethics

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

References

- Adams, L. C., & Macdonald, S. E. (2018). Spontaneous preference for primate photographs in Sumatran orangutans (*Pongo abelii*). *International Journal of Comparative Psychology*, *31*, 1–16. <https://doi.org/10.46867/ijcp.2018.31.04.05>
- Albuquerque, N., Guo, K., Wilkinson, A., Savalli, C., Otta, E., & Mills, D. (2016). Dogs recognize dog and human emotions. *Biology Letters*, *12*(1), 20150883. <https://doi.org/10.1007/s10071-010-0316-y>
- Almeling, L., Hammerschmidt, K., Sennhenn-Reulen, H., Freund, A. M., & Fischer, J. (2016). Motivational shifts in aging monkeys and the origins of social selectivity. *Current Biology*, *26*(13), 1744–1749. <https://doi.org/10.1016/j.cub.2016.04.066>
- Aslin, R. N. (2007). What’s in a look? *Developmental Science*, *10*(1), 48–53. <https://doi.org/10.1111/j.1467-7687.2007.00563.x>
- Bellegarde, L. G. A., Haskell, M. J., Duvaux-Ponter, C., Weiss, A., Boissy, A., & Erhard, H. W. (2017). Face-based perception of emotions in dairy goats. *Applied Animal Behaviour Science*, *193*, 51–59. <https://doi.org/10.1016/j.applanim.2017.03.014>
- Bethell, E. J., Holmes, A., MacLarnon, A., & Semple, S. (2012). Evidence that emotion mediates social attention in rhesus macaques. *PLoS ONE*, *7*(8). <https://doi.org/10.1371/journal.pone.0044387>
- Bird, C. D., & Emery, N. J. (2008). Using video playback to investigate the social preferences of rooks, *Corvus*

- frugilegus*. *Animal Behaviour*, 76(3), 679–687. <https://doi.org/10.1016/j.anbehav.2008.04.014>
- Bolker, B. (2019). Linear Mixed-Effects Models using “Eigen” and S4.
- Calapai, A., Berger, M., Niessing, M., Heisig, K., Brockhausen, R., Treue, S., & Gail, A. (2016). A cage-based training, cognitive testing and enrichment system optimized for rhesus macaques in neuroscience research. *Behavior Research Methods*, 49(1), 35–45. <https://doi.org/10.3758/s13428-016-0707-3>
- Clark, D. L., & Uetz, G. W. (1993). Signal efficacy and the evolution of male dimorphism in the jumping spider, *Maevia inclemens*. *Proceedings of the National Academy of Sciences of the United States of America*, 90(24), 11954–11957. <https://doi.org/10.1073/pnas.90.24.11954>
- Costa, M., Gomez, A., Barat, E., Lio, G., Duhamel, J.-R., & Sirigu, A. (2018). Implicit preference for human trustworthy faces in macaque monkeys. *Nature Communications*, 9(1), 4529. <https://doi.org/10.1038/s41467-018-06987-4>
- D’Eath, R. B. (1998). Can video images imitate real stimuli in animal behaviour experiments? *Biological Reviews*, 73(3), 267–292. <https://doi.org/10.1111/j.1469-185X.1998.tb00031.x>
- Dubuc, C., Allen, W. L., Maestriperi, D., & Higham, J. P. (2014). Is male rhesus macaque red color ornamentation attractive to females? *Behavioral Ecology and Sociobiology*, 68(7), 1215–1224. <https://doi.org/10.1007/s00265-014-1732-9>
- Dufour, V., Pascalis, O., & Petit, O. (2006). Face processing limitation to own species in primates: A comparative study in brown capuchins, Tonkean macaques and humans. *Behavioural Processes*, 73(1), 107–113. <https://doi.org/10.1016/j.beproc.2006.04.006>
- Fagot, J. (2000). *Picture perception in animals*. Psychology Press Ltd.
- Fagot, J., Thompson, R. K. R., & Parron, C. (2010). How to read a picture: Lessons from nonhuman primates. *PNAS*, 107(2), 519–520. <https://doi.org/10.1073/pnas.0913577107>
- Fantz, R. L. (1958). Pattern vision in young infants. *The Psychological Record*, 8, 43–47.
- Farroni, T., Menon, E., & Johnson, M. H. (2006). Factors influencing newborns’ preference for faces with eye contact. *Journal of Experimental Child Psychology*, 95(4), 298–308. <https://doi.org/10.1016/j.jecp.2006.08.001>
- Fujita, K. (1987). Species recognition by five macaque monkeys. *Primates*, 28(3), 353–366. <https://doi.org/10.1007/BF02381018>
- Gothard, K. M., Erickson, C. A., & Amaral, D. G. (2004). How do rhesus monkeys (*Macaca mulatta*) scan faces in a visual paired comparison task? *Animal Cognition*, 7(1), 25–36. <https://doi.org/10.1007/s10071-003-0179-6>
- Hopper, L. M., Gulli, R. A., Howard, L. H., Kano, F., Krupenye, C., Ryan, A. M., & Paukner, A. (2020). The application of noninvasive, restraint-free eye-tracking methods for use with nonhuman primates. *Behavior Research Methods*. Advance online publication. <https://doi.org/10.3758/s13428-020-01465-6>
- Kano, F., Call, J., & Tomonaga, M. (2012). Face and eye scanning in gorillas (*Gorilla gorilla*), orangutans (*Pongo abelii*), and humans (*Homo sapiens*): Unique eye-viewing patterns in humans among hominids. *Journal of Comparative Psychology*, 126(4), 388–398. <https://doi.org/10.1037/a0029615>
- Kim, Y., Choi, Y.-K., & Emery, S. (2013). Logistic regression with multiple random effects: A simulation study of estimation methods and statistical packages. *Journal of the American Statistical Association*, 67(3). <https://doi.org/doi:10.1080/00031305.2013.817357>
- Koba, R., & Izumi, A. (2008). Japanese monkeys (*Macaca fuscata*) discriminate between pictures of conspecific males and females without specific training. *Behavioural Processes*, 79(1), 70–73. <https://doi.org/10.1016/j.beproc.2008.04.005>
- Kohda, M., Jordan, L. A., Hotta, T., Kosaka, N., Karino, K., Tanaka, H., Taniyama, M., & Takeyama, T. (2015). Facial recognition in a group-living cichlid fish. *PLoS ONE*, 10(11), e0142552. <https://doi.org/10.1371/journal.pone.0142552>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, 354(6308), 110–114. <https://doi.org/10.1126/science.aaf8110>
- Lyons, M., Marcinkowska, U., Moisey, V., & Harrison, N. (2016). The effects of resource availability and relationship status on women’s preference for facial masculinity in men: An eye-tracking study. *Personality and Individual Differences*, 95, 25–28. <https://doi.org/10.1016/j.paid.2016.02.025>
- Méary, D., Wu, L., Zhinan, L., Guo, K., & Pascalis, O. (2014). Seeing two faces together: Preference formation in humans and rhesus macaques. *Animal Cognition*, 17(5), 1107–1119. <https://doi.org/10.1007/s10071-014-0742-3>

- Morton, F. B., Brosnan, S. F., Prétôt, L., Buchanan-Smith, H. M., O'Sullivan, E., Stocker, M., D'Mello, D., & Wilson, V. A. D. (2016). Using photographs to study animal social cognition and behaviour: Do capuchins' responses to photos reflect reality? *Behavioural Processes*, *124*, 38–46. <https://doi.org/10.1016/j.beproc.2015.10.005>
- Nawroth, C., Albuquerque, N., Savalli, C., Single, M., & McElligott, A. G. (2018). Goats prefer positive human emotional facial expressions. *Royal Society Open Science*, *5*(8), 180491. <https://doi.org/10.1098/rsos.180491>
- Overduin-de Vries, A. M., Bakker, F. A. A., Spruijt, B. M., & Sterck, E. H. M. (2016). Male long-tailed macaques (*Macaca fascicularis*) understand the target of facial threat. *American Journal of Primatology*, *78*(7), 720–730. <https://doi.org/10.1002/ajp.22536>
- Paukner, A., Wooddell, L. J., Lefevre, C. E., Lonsdorf, E., & Lonsdorf, E. (2017). Do capuchin monkeys (*Sapajus apella*) prefer symmetrical face shapes? *Journal of Comparative Psychology*, *131*(1), 73–77. <https://doi.org/10.1037/com0000052>
- Pfefferle, D., Kazem, A. J. N., Brockhausen, R. R., Ruiz-Lambides, A. V., & Widdig, A. (2014). Monkeys spontaneously discriminate their unfamiliar paternal kin under natural conditions using facial cues. *Current Biology*, *24*(15), 1806–1810. <https://doi.org/10.1016/j.cub.2014.06.058>
- Rakotonirina, H., Kappeler, P. M., & Fichtel, C. (2018). The role of facial pattern variation for species recognition in red-fronted lemurs (*Eulemur rufifrons*). *BMC Evolutionary Biology*, *18*(1), 1–10. <https://doi.org/10.1186/s12862-018-1126-0>
- Sato, A., Koda, H., Lemasson, A., Nagumo, S., & Masataka, N. (2012). Visual recognition of age class and preference for infantile features: Implications for species-specific vs universal cognitive traits in primates. *PLoS ONE*, *7*(5), e38387. <https://doi.org/10.1371/journal.pone.0038387>
- Satoh, S., Tanaka, H., & Kohda, M. (2016). Facial recognition in a Discus fish (Cichlidae): Experimental approach using digital models. *PLoS One*, *11*(5), 1–11. <https://doi.org/10.1371/journal.pone.0154543>
- Schell, A., Rieck, K., Schell, K., Hammerschmidt, K., & Fischer, J. (2011). Adult but not juvenile Barbary macaques spontaneously recognize group members from pictures. *Animal Cognition*, *14*(4), 503–509. <https://doi.org/10.1007/s10071-011-0383-8>
- Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317–1322. <https://doi.org/10.1038/nn1150>
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. <https://doi.org/10.1037/0033-2909.86.2.420>
- Somppi, S., Törnqvist, H., Hänninen, L., Krause, C., & Vainio, O. (2012). Dogs do look at images: Eye tracking in canine cognition research. *Animal Cognition*, *15*(2), 163–174. <https://doi.org/10.1007/s10071-011-0442-1>
- von Rohr, C., Van Schaik, C. P., Kissling, A., & Burkart, J. M. (2015). Chimpanzees' bystander reactions to infanticide an evolutionary precursor of social norms? *Human Nature*, *26*(2), 143–160. <https://doi.org/10.1007/s12110-015-9228-5>
- Waïtt, C., & Little, A. C. (2006). Preferences for symmetry in conspecific facial shape among *Macaca mulatta*. *International Journal of Primatology*, *27*(1), 133–145. <https://doi.org/10.1007/s10764-005-9015-y>
- Waïtt, C., Little, A. C., Wolfensohn, S., Honess, P., Brown, A. P., Buchanan-Smith, H. M., & Perrett, D. I. (2003). Evidence from rhesus macaques suggests that male coloration plays a role in female primate mate choice. *Proceedings of the Royal Society B: Biological Sciences*, *270*(suppl_2), S144–S146. <https://doi.org/10.1098/rsbl.2003.0065>
- Wathan, J., Proops, L., Grounds, K., & McComb, K. (2016). Horses discriminate between facial expressions of conspecifics. *Scientific Reports*, *6*(38322), 1–11. <https://doi.org/10.1038/srep38322>
- Wilson, V. A. D., Altschul, D. M. & Weiss, A. (in prep). Looking time and video scene continuity as measures of emotional perspective taking in chimpanzees.
- Wilson, V. A. D., Kade, C., Moeller, S., Treue, S., Kagan, I., & Fischer, J. (2020). Macaque gaze responses to the Primatar: A virtual macaque head for social cognition research. *Frontiers in Psychology*, *11*, 1–13. <https://doi.org/10.3389/fpsyg.2020.01645>
- Winters, S., Dubuc, C., & Higham, J. P. (2015). Perspectives: The looking time experimental paradigm in studies of animal visual perception and cognition. *Ethology*, *121*(7), 625–640. <https://doi.org/10.1111/eth.12378>
- Zhang, H., Lu, N., Feng, C., Thurston, S. W., Xia, Y., & Tu, X. M. (2011). On fitting generalized linear mixed-effects models for binary responses using different statistical packages. *Statistics in Medicine*, *30*(20), 2562–2572. <https://doi.org/10.1002/sim.4265>

Appendix

Experiment 1

Viewing Task 1

Apparatus for Training. For viewing task 1, 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 the 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 in. 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 1 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 2, where monkeys only saw each image once.

Choice Task 1

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 these data to analyze 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 the monkeys' short attention span, this was a challenging task for them 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 s instead of 300 s; 160 trials when the enlarged image was displayed for 10 instead of 6 s; 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 Figure A1). 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]; object: $OR = 1.40$, $SE = 0.08$, $p < .001$, 95% CI [1.21, 1.64]; see Figure A1). These findings are consistent with the full model, assessing response across all trials.

Experiment 2

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 color 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 to 3 years old) or subadult (3 to 4.5 years old) 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 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 ($SD = 0.90$).

Sensitivity Analyses

Viewing Task 2. For viewing task 2, 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 (Figure A2).

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 Figure A3).

Finally, we assessed, 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 2. 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 Table A2).

References

- de Waal, F. B., van Hooff, J. A., & Netto, W. J. (1976). An ethological analysis of types of agonistic interaction in a captive group of Java-monkeys (*Macaca fascicularis*). *Primates*, *17*(3), 257-290.
<https://doi.org/10.1007/BF02382785>
- Fujita, K. (1987). Species recognition by five macaque monkeys. *Primates*, *28*(3), 353–366.
<https://doi.org/10.1007/BF02381018>
- 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>

Table A1*Demographics and Participation of Study Participants Across Tasks*

ID	Age at start of testing (yrs) ¹	Sex	VT1 ELO training Training	VT1 XBI		Choice		VT2 Test	Free view Test
				Training	Test	Training	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 2 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.

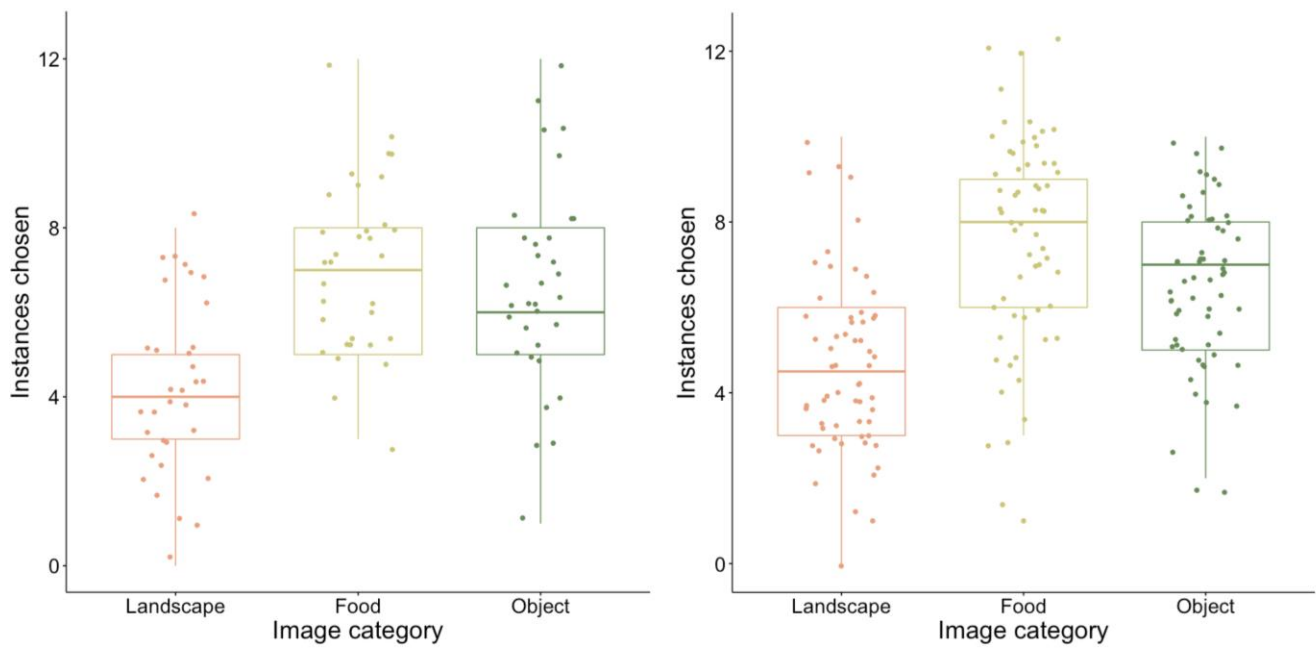
Table A2*Total Number of Instances of Image Choice and First Touch from Choice Task 2*

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).

Figure A1

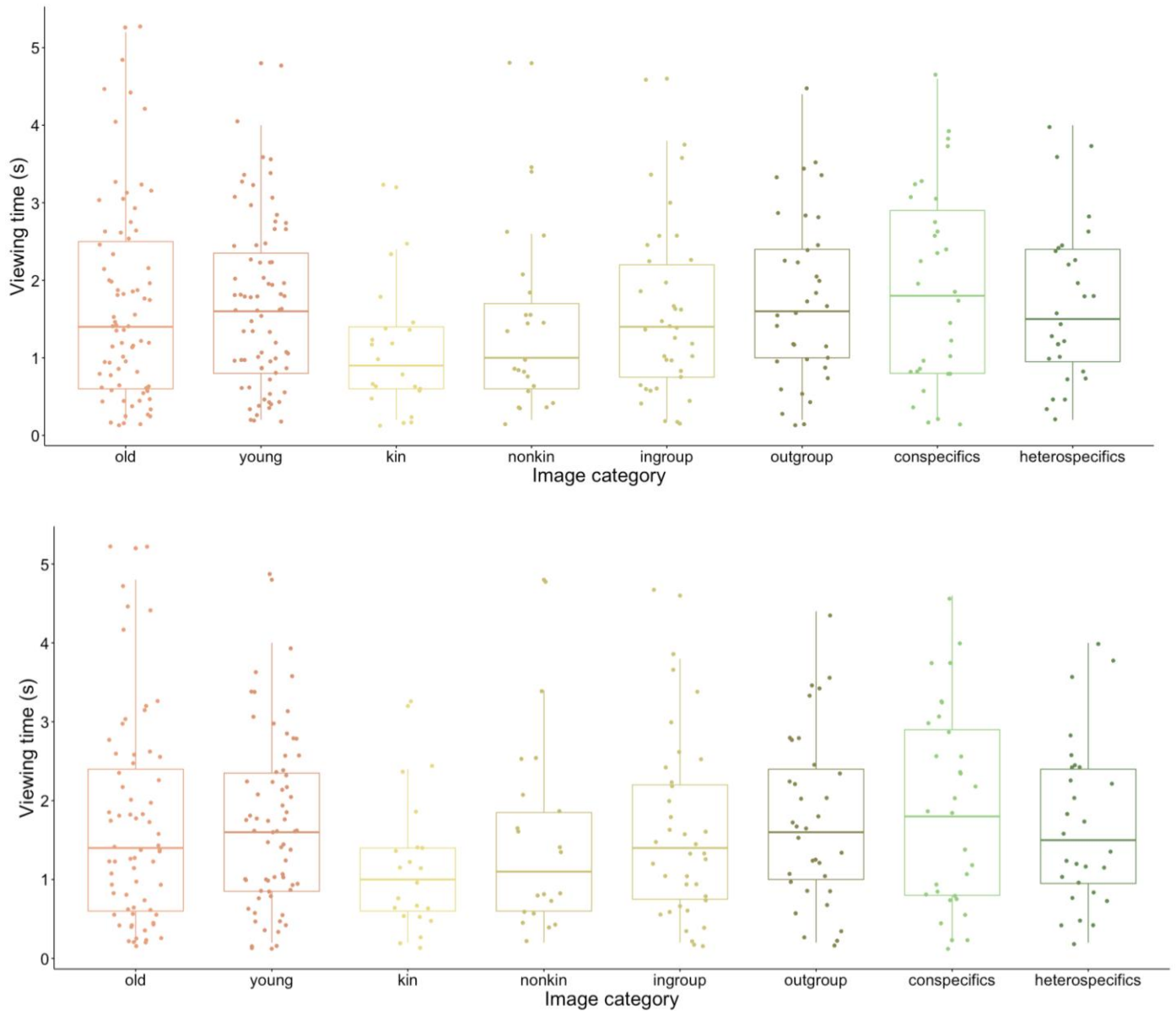
Image Preferences on Choice Task 1, Before (Left) and After (Right) Procedural Changes



Note. 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 A2

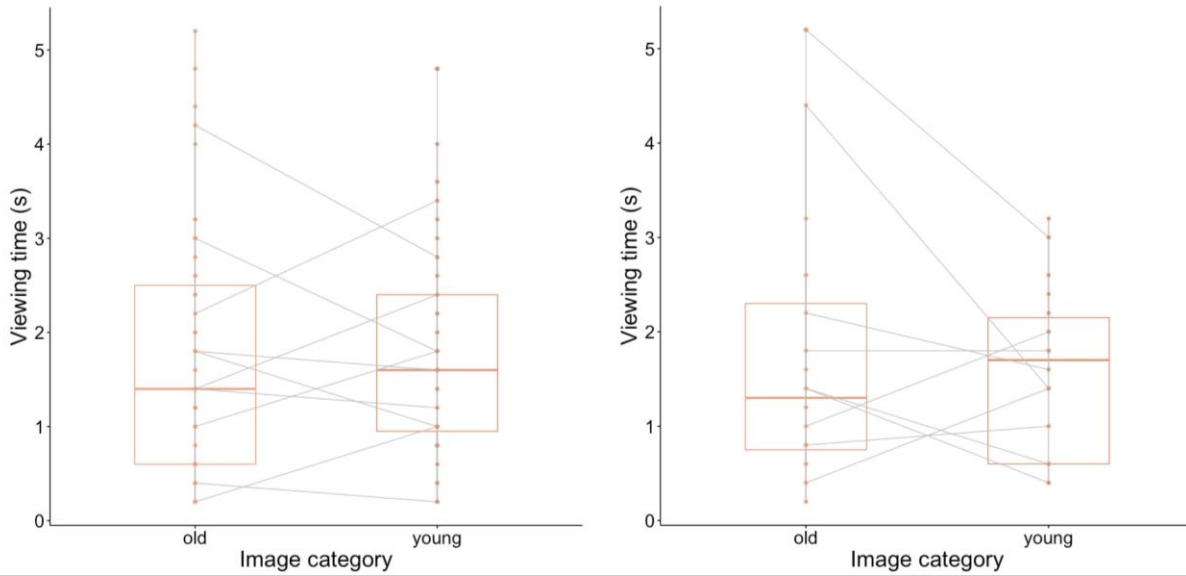
Looking Time to Social Image Categories on Viewing Task 2, For All Images (Top) and With Mirror-Flipped Images Removed (Bottom)



Note. 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 A3

Looking Time on Viewing Task 2 to Age-Category Image Pairs with Different Age Gaps



Note. 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.