Author Manuscript

Raisa Hernández-Pacheco ORCID iD: 0000-0002-3681-5127
Alexandra Rosati ORCID iD: 0000-0002-6906-7807

Individual differences in sociocognitive traits in semi-free-ranging rhesus monkeys *(Macaca mulatta)*

Alexis A. Diaz[1,2]*, Raisa Hernández-Pacheco[2], & Alexandra G. Rosati[3]*
[1]Stanford University, Department of Biology
[2]California State University, Long Beach, Department of Biological Sciences
[3]University of Michigan, Departments of Psychology and Anthropology
*Correspondence to rosati@umich.edu or alexisdz@stanford.edu

Abstract

Characterizing individual differences in cognition is crucial for understanding the evolution of cognition as well as to test the biological consequences of different cognitive traits. Here, we harnessed the strengths of a uniquely large, naturally-living primate population at the Cayo Santiago Biological Field Station to characterized individual differences in rhesus monkey performance across two social cognitive tasks. A total of n = 204 semi-free-ranging adult rhesus monkeys participated in a data collection procedure, where we aimed to test individuals on both tasks at two time-points one year apart. In the *socioemotional responses task*, we assessed monkeys' attention to conspecific photographs with neutral versus negative emotional expressions. We found that monkeys showed overall declines in interest in conspecific photographs with age, but relative increases in attention to threat stimuli specifically, and further that these responses exhibited long-term stability across repeated testing. In the *gaze following task* we assessed monkeys' propensity to co-orient with an experimenter. Here, we found evidence for age-related change in responses, and responses show only limited evidence for repeatability over time. Finally, we found some evidence for common individual variation for performance across the tasks: monkeys that showed greater interest in conspecific photographs were more

This is the author manuscript accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ajp.23660.

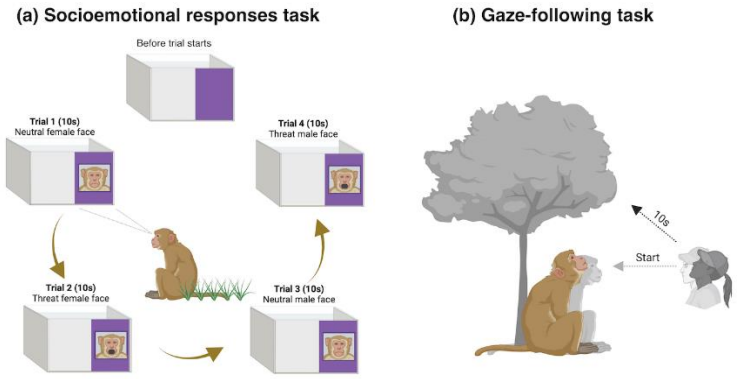
This article is protected by copyright. All rights reserved.

Author Manuscript

likely to follow a human's gaze. These results show how studies of comparative cognitive development and aging can provide insights into the evolution of cognition, and identify core primate social cognitive traits that may be related across and within individuals.

Graphical Abstract

We tested a large sample of monkeys longitudinally one two sociocognitive tasks to examine individual differences in longitudinal change in cognitive responses. (a) In the socioemotional responses task, we measured preferences for viewing photographs of conspecifics producing emotional versus neutral expressions. (b) In the gaze following task, we measured co-orienting responses with a demonstrator. We found that monkeys showed repeatable responses with clear age-related differences in the socioemotional responses task, but not in the gaze following task. In addition, monkeys that showed greater interest in conspecific photographs were more likely to follow a human's gaze.



Research Highlights

- We examined individual variation and longitudinal patterns of cognitive performance in semi-free-ranging rhesus macaques, using two social cognitive tasks
- Monkeys showed distinct age-related change in a task measuring responses to socioemotional signals, and their performance exhibited long-term repeatability
- Monkeys did not show clear age shifts in a task measuring gaze following (co-orienting) responses, and individual performance was not clearly stable over time

This article is protected by copyright. All rights reserved.

Author Manuscript

- Individual differences approaches can provide new insights into the evolution of cognitive abilities across species

Key words: primate cognition, comparative development, emotional processing, gaze following

Abbreviations*:* None

1. Introduction

Comparative studies of cognition in animal species are important for understanding the evolutionary history and biological consequences of different cognitive skills (Pritchard, Hurly, Tello-Ramos, & Healy, 2016; Rosati, Machanda, & Slocombe, 2022; Thornton & Truskanov, 2022). While much work on animal cognition has focused on testing whether different species have a particular skill and in comparing variation in performance across species (Bond, Kamil, & Balda, 2007; De Petrillo, Bettle, & Rosati, 2022; Herrmann, Call, Hernadez-Lloreda, Hare, & Tomasello, 2007; Maclean et al., 2014), there is increasing appreciation that an individual differences approach to examining intraspecific differences—how and why individuals of the same population or species solve problems—is also crucial for understanding how and why cognition evolves (Thornton & Lukas, 2012; Völter, Tinklenberg, Call, & Seed, 2018). In particular, an individual differences approach aimed at characterizing intraspecific variation in cognition is needed to address several core questions concerning the evolutionary function of cognition and behavior, including whether animals show age-related change in cognition as they develop; whether sexes differ in their responses to problems; and finally, whether cognitive performance represents a stable and potentially heritable trait.

This article is protected by copyright. All rights reserved.

Author Manuscript

First, an individual differences approach allows for characterization of age-related changes in cognition that shape how different individuals respond to important challenges in their environment (Rathke & Fischer, 2020; Rosati, Arre, Platt, & Santos, 2016; Soravia, Ashton, Thornton, & Ridley, 2022). It is well-recognized that human cognition develops and changes over the lifespan, and emerging work on comparative cognitive development aims to compare the ontogeny of different cognitive abilities across the life course in different species (Gomez, 2005; Matsuzawa, 2007; Matsuzawa, Tomonaga, & Tanaka, 2006; Rosati, Wobber, Hughes, & Santos, 2014). Integrating perspectives from both developmental psychology and evolutionary biology is crucial to understand how animals shift their cognition and behavior during different life stages (Almeling, Hammerschmidt, Senn-Reulen, Freund, & Fischer, 2016; Almeling, Senn-Reulen, Hammerschmidt, Freund, & Fischer, 2017; Fischer, 2017b; Haux, Engelmann, Arslan, Hertwug, & Herrmann, 2022; Machanda & Rosati, 2020; Rathke & Fischer, 2020; Rosati et al., 2016; Rosati, Arre, Platt, & Santos, 2018; Rosati, Emery Thompson, Atencia, & Buckholtz, 2023; Rosati et al., 2020). For example, species can vary widely in their life history characteristics compared to humans (Alberts et al., 2013), and differences in the pace of growth and length of lifespan has been linked to important differences in brain maturation during the juvenile period and age-related senescence (Leigh, 2004, 2012). As such, studies of comparative development can test how cognitive shifts across the lifespan enable adaptive behavior at different life history phases as well as how these shifts are connected to physiological changes in the mind and body.

Another important aspect of individual differences in cognition concerns whether there are persistent sex differences in performance. While many species show pervasive

This article is protected by copyright. All rights reserved.

Author Manuscript

sex differences in their adult behavior, surprisingly there are few demonstrations of similar differences in cognition. For example, chimpanzees are an extremely well-studied primate both in terms of behavior and cognition, and show pervasive differences in patterns of social behavior and foraging that are thought to reflect adaptive differences in how males versus females maximize their reproductive success (Boesch & Boesch, 1981; Gilby et al., 2017; Lonsdorf, 2005; Machanda, Gilby, & Wrangham, 2013; Sabbi et al., 2021; Thompson Gonzalez et al., 2021). For example, males are much more social than females in wild populations (Gilby & Wrangham, 2008; Machanda et al., 2013; Pusey, Pintea, Wilson, Kamenya, & Goodall, 2007; Wrangham, 2000). One possibility is that these differences in behavior reflect fundamental differences in underlying cognitive mechanisms—for example, males may have more robust social cognition. However, current work suggests few sex differences in a variety of cognitive responses in chimpanzees, including in social cognition (De Petrillo & Rosati, 2021; Herrmann, Hare, Call, & Tomasello, 2010; Wobber, Herrmann, Hare, Wrangham, & Tomasello, 2014) (but see Cantwell, Buckholtz, Atencia, & Rosati, 2022; Haux et al., 2022). Yet it is also important to note that many studies of primate cognition are limited by small sample sizes, highlighting the importance of examining large samples of animals to test both sex and age-related shifts.

Characterizing individual differences is further a crucial step for testing what components of cognition are stable, potentially heritable, and thus impact fitness outcomes (Ashton, Ridley, Edwards, & Thornton, 2018; Bohn et al., 2023; Cauchoix, 2018; Croston, Branch, Kozlovsky, Dukas, & Pravosudov, 2015; Pravosudov, 2022). A primary question here is whether performance on cognitive tasks reflects some stable trait. This is important

This article is protected by copyright. All rights reserved.

Author Manuscript

because stable differences (repeatability) in behaviors sets the upper limit for the heritability of this behavior (Cauchoix, 2018; Wilson et al., 2010). Heritability is a crucial prerequisite for natural selection to act on the trait, to and more generally for cognition to have some consistent impact on individual's biological outcomes such as health and longevity. Experimental measures are often assumed to reflect such a stable individual propensity, indexing the skillfulness with which a given individual solve particular problems, yet there is also accumulating evidence that many fleeting contextual factors can influence task performance (Schubiger, Fichtel, & Burkart, 2020). For example, animals may differ in their motivation, hunger, or responses to novel aspects of the task (De Petrillo, Nair, Cantwell, & Rosati, 2022; Herrmann, Hare, Cissewski, & Tomasello, 2011)—all factors that partially reflect transient aspects of their environment as opposed to stable differences. Thus, even though much work in primate cognition has tested animals at a single time point, approaches that assess the repeatability of cognitive performance over time (e.g., Ashton et al., 2018; Bohn et al., 2023; Cauchoix, 2018) are crucial to understand these patterns. While there is increasing evidence for repeatability of behavioral patterns in non-human primates (Brent, Semple, Maclarnon, Ruiz-La, & Platt, 2014; Tkaczynski et al., 2020), most work to date on repeatability of cognitive performance in animals has examined non-primates (see Cauchoix, 2018 for a metanalysis of studies) (but see Bohn et al., 2023 for evidence from great ape cognition). As such, little is currently known about repeatability of primate cognitive performance.

A final set of questions that can be addressed by an individual differences approach is whether or how performance across different tasks is related. Recently there has been increasing interest in testing the same animals on a larger battery of tasks to characterize

This article is protected by copyright. All rights reserved.

Author Manuscript

this kind of individual variation in animal cognitive abilities (Shaw & Schmelz, 2017). For example, one battery of multiple cognitive tasks spanning multiple domains has now been implanted with multiple primate species including lemurs, monkeys, nonhuman apes, and children (Fichtel, Dinter, & Kappeler, 2020; Herrmann et al., 2007; Herrmann, Hare, et al., 2010; Herrmann, Hernández-Lloreda, Call, Hare, & Tomasello, 2010; Schmitte, Pankau, & Fischer, 2012). Several other studies to date have used two or more tasks to examine interrelationships between different skills (Beran & Hopkins, 2018; De Petrillo, Nair, et al., 2022; Rosati, DiNicola, & Buckholtz, 2018; Völter et al., 2022; Völter et al., 2018). From a cognitive perspective, this approach can elucidate the underlying structure of cognition, by assessing if performance on different behavioral tasks reflects a shared versus divergent underlying cognitive structure in terms of the cognitive processes supporting performance. Similarly, from a biological perspective this can give new insights into trait covariation.

In the current study, we used an individual difference approach to examine patterns of social cognition in a large population of semi-free-ranging rhesus macaques (*Macaca mulatta*). We focused on social cognitive performance in this particular species for several reasons. First, rhesus monkeys are a highly social primate species that live in large mixed-sex groups with a variety of social interactions between both kin and non-kin (Ellis, Snyder-Mackler, Ruiz-Lambides, Platt, & Brent, 2019; Rawlins & Kessler, 1986), and as such social cognitive skills are thought to be important for adaptive behavior in this socially-complex species (Drayton & Santos, 2015; Ghazanfar & Santos, 2004). Second, rhesus share many human-relevant cognitive features and in general social cognition is fairly well-characterized in this species. For example, macaques share human-like social

This article is protected by copyright. All rights reserved.

Author Manuscript

preferences for attending to socially relevant information including gaze direction, emotional expressions in others, and other's perceptual and goal states (Drayton & Santos, 2015; Drayton & Santos, 2017; Ghazanfar & Santos, 2004; Gothard, Battaglia, Erickson, Spitler, & Amaral, 2007; Hoffman, Gothard, Schmid, & Logothetis, 2007; Rosati et al., 2016; Rosati, Arre, et al., 2018; Santos, Nissen, & Ferrugia, 2006). Finally**,** while rhesus macaques have a fairly long-life span for a mammal, they also have a shorter lifespan than humans (Alberts et al., 2013; Bogin & Smith, 1996; Leigh, 2004, 2012; Tigges, Gordon, McClure, Hall, & Peters, 1988). As such, they may exhibit important differences in patterns of cognitive development and aging compared to humans. Indeed, work to date has revealed a mixed picture concerning shared versus divergent patterns of cognitive development, with shared patterns in some contexts but not others (Hernández-Pacheco, Steiner, Rosati, & Tuljapurka, 2023; Newman et al., 2023)

We specifically focused on two distinct aspects of social cognition in rhesus macaques. First, we tested animals' responses to socioemotional signals by measuring their interest and attention to conspecific photographs. In humans, there are well-documented changes in socioemotional processing across the lifespan. For example, whereas children and younger adults exhibit a 'negativity' bias, selectively attending to and remembering negative stimuli such as negative emotional faces, older adults rather exhibit a positivity bias showing less interest in such faces (Carstensen & Mikels, 2005; Carstensen et al., 2011; Mather & Carstensen, 2005). Understanding these kinds of shifts during aging has also emerged as an important target of comparative research on other animals (Almeling et al., 2016; Almeling et al., 2017; Fischer, 2017a; Machanda & Rosati, 2020; Rosati et al., 2020; Thompson Gonzalez et al., 2021). In particular, some experimental work with cross-

This article is protected by copyright. All rights reserved.

Author Manuscript

sectional sample of rhesus monkeys ranging from juvenility to old age comparing their duration of looking towards photos of negative conspecific signals (threat displays) versus matched neutral expressions revealed that while monkeys show clear age-related shifts in responses, it was actually the opposite pattern of that observed in humans: older monkeys were more attentive to negative emotional stimuli than were younger monkeys (Rosati, Arre, et al., 2018).

Second, we tested monkeys' abilities to gaze follow, or co-orient with others. In humans, gaze following is a foundational social skill that emerges early in infancy (D'Entremont, Hains, & Muir, 1997), and is thought to be a crucial building block for later-emerging social cognitive skills, including both communication and theory of mind skills such as inferring other's perceptions (Bettle & Rosati, 2021; Brooks & Meltzoff, 2005; Emery, 2000; Shepherd, 2010; Wellman, 2011). Yet gaze following is also a good target for studies of comparative development as it shows characteristic shifts across the lifespan in humans: adult women exhibit greater responsivity to gaze cues than do men (Alwall, Johansson, & Hansen, 2010; Bayliss, di Pellegrino, & Tipper, 2005; Deaner, Shepherd, & Platt, 2007; Mundy et al., 2007), and older adults show declines in responsivity to gaze cues (Kuhn, Pagano, Maani, & Bunce, 2014; Slessor, Phillips, & Bull, 2008; Slessor et al., 2014). Many nonhuman animals also follow gaze to some degree (Rosati & Hare, 2009; Shepherd, 2010), and there is also emerging evidence concerning the developmental trajectory of this skill across species: some primate species appear to develop gaze following skills much more slowly than humans (Ferrari, Coude, Gallese, & Fogassi, 2008; Ferrari, Kohler, Fogassi, & Gallese, 2000; Okamoto et al., 2002; Okamoto-Barth, Tomonaga, Tanaka, & Matsuzawa, 2008; Teufel, Gutmann, Pirow, & Fischer, 2010;

This article is protected by copyright. All rights reserved.

Author Manuscript

Tomasello & Carpenter, 2005; Tomasello, Hare, & Fogleman, 2001; Tomonaga, Tanaka, Matsuzawa, Myowa-Yamakoshi, & Kosugi, 2004; Wobber et al., 2014). Conversely, prior work with rhesus monkeys specifically suggests they show a pattern mirroring humans, with this skill emerging in infancy and then declining in old age (Rosati et al., 2016; Tomasello et al., 2001).

In the current study, we adapted methods from prior work examining socioemotional biases (Rosati, Arre, et al., 2018) and gaze following (Rosati et al., 2016) in the Cayo Santiago rhesus macaques using pre-registered data collection methods. While those prior studies used a cross-sectional approach testing monkeys from juvenility to old age, here we rather used these validated tasks to collect systematic longitudinal data from a large sample of monkeys on both tasks. Specifically, we tested each monkey on both tasks at two different time points, approximately one year apart to capture long-term patterns in cognitive performance. In addition, our sample here consisted of primarily adults to focus on individual variation and age-related shifts during the adult lifespan. This approach allowed us to test several interrelated questions about individual differences in cognition: (1) whether individuals show any consistent age- or sex-related differences in cognitive responses; (2) whether performance in tasks reflects traits that are stable over time; and (3) whether an individual's performance across these different cognitive metrics is related.

2. Methods

In our pre-registered data collection procedure (https://aspredicted.org/48K_8XP), we tested a sample of adult monkeys from the Cayo Santiago Biological Field Station on two sociocognitive tasks (see Figure 1). The *socioemotional responses task* measured

This article is protected by copyright. All rights reserved.

Author Manuscript

preferences for viewing photographs of conspecifics producing emotional versus neutral expressions, and the *gaze following task* measured co-orienting responses with a demonstrator. As feasible, individual animals were tested in both tasks at two timepoints, approximately one year apart.

2.1 Ethics statement

All non-invasive behavioral tests reported in this paper were approved by the Institutional Animal Care and Use Committee of the University of Puerto Rico Medical Sciences Campus (protocol #A140116) and adhered to site guidelines for animal research. Research procedures adhered to the American Society of Primatologists Principles for the Ethical Treatment of Nonhuman Primates.

2.2 Subjects

Our subjects were rhesus monkeys from the Cayo Santiago Biological Field Station in Puerto Rico. The Cayo Santiago population consists of over 1500 individually identifiable monkeys living in multiple naturally formed social groups on a 38 acre island off the coast of Puerto Rico (Rawlins & Kessler, 1986). Animals are well-habituated to human observers and are identifiable based on a combination of unique tattoos (three-character codes), ear notches, as well as aspects of physical appearance. Many monkeys have participated in prior cognitive studies, including tasks involving similar procedures as the current work (Bettle & Rosati, 2019; Dubuc et al., 2016; Higham et al., 2011; Rosati et al., 2016; Rosati, Arre, et al., 2018). For the purposes of our study, we had a target set of primarily adult animals as we were interested in variation in the adult lifespan, rather than early life development. We also focused on monkeys living in two of the groups at Cayo Santiago (F and V) to integrate this cognitive data with other information collected

This article is protected by copyright. All rights reserved.

Author Manuscript

on animals in those groups, such as aspects of their dominance rank and social behavior (these analyses are not part of the current work).

We tested a final sample 204 monkeys that completed at least one task across the two years of testing. This included 101 females and 103 males, with a mean age of 12.0 years (range: 4.7-29.7 years). In this free-ranging context, not all individuals could be successfully tested in both tasks in both years (see Table 1 for breakdown of participation across tasks and years). Across analyses, we therefore use different subsets of these monkeys as relevant to test different questions.

2.3 Field experimental methods overview

In the general testing procedure for both tasks, two experimenters walked around the island searching for target monkeys that were sitting calmly and located in an appropriate position for the experimenters to test them (e.g., in a location where we could set the apparatus for the socioemotional task, and near an appropriate tree for the gaze task as noted below). On a given day, each monkey would complete one session of one task; assuming they were located again, they would complete the second task at a later point, typically on a different day when they were similarly located in an appropriate testing position. For monkeys that completed both tasks in a given year, these tests occurred within an approximately two-month window of testing. Because of the nature of cognitive testing procedures in this free-ranging context, we did not systematically approach animals for the two tests in any particular order, but rather tested them as feasible when they were located and individually identified depending on which study equipment we had at the time. Then one year later, we conducted the same tests again, attempting to locate the same set of target monkeys within the two-month testing window.

This article is protected by copyright. All rights reserved.

Author Manuscript

In a given testing session, experimenter one, the demonstrator, positioned herself approximately 2-3m away from the monkey. In the socioemotional responses task, she knelt in this location and placed an apparatus on the ground in front of her, so that the monkey could observe successive photos during the task (see Figure 1a and details below). In the gaze following task, she stood in front of the monkey from this distance, and produced gaze cues (see Figure 1b and details below). During both tasks, experimenter two knelt or stood behind the presenter, and filmed the monkey's face with a camera so that their responses could be coded from video.

2.4 Socioemotional responses task procedures

This task aimed to characterize individual variation in responses to conspecific photographs that varied in expression (neutral versus threat faces), using procedures following those previously implemented with this population (Rosati, Arre, et al., 2018). As in that prior work, monkeys could view a sequence of photos of unfamiliar conspecifics producing neutral or emotional facial expression, and we measured their looking time to each photo to index relative interest and motivation to view these various stimuli. The first set of two trials showed photos of a female producing a neutral expression, followed by the same female producing a threat expression. The second set of two trials showed a male producing a neutral expression followed by the same male producing a threat expression.

This article is protected by copyright. All rights reserved.

Author Manuscript

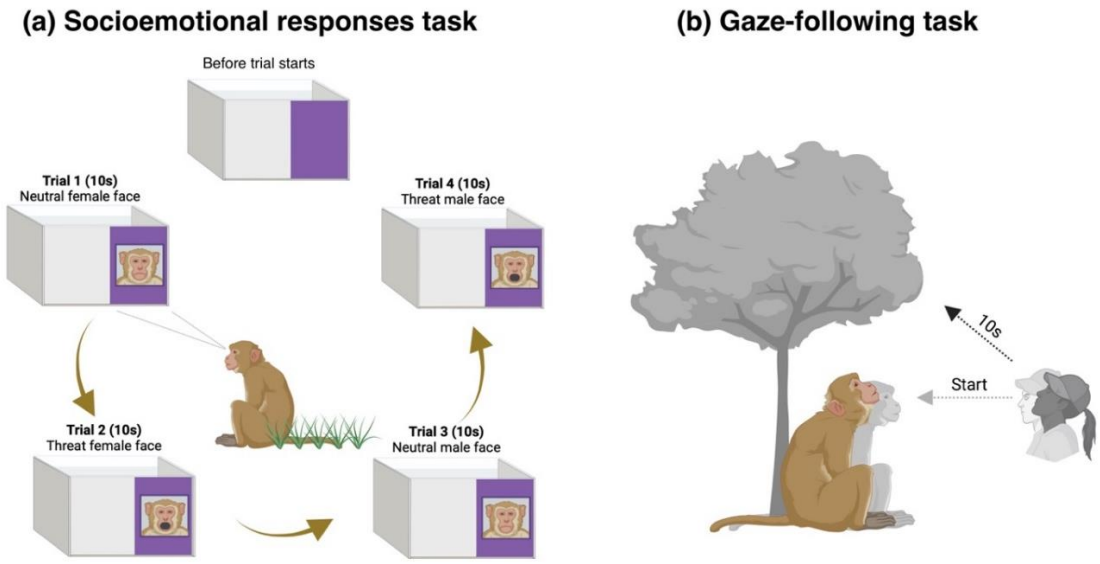


Figure 1. Experimental setup for the sociocognitive tasks. (a) The socioemotional responses task measured preferences for viewing photographs of conspecifics producing emotional versus neutral expressions. Monkeys were presented with up to four photos, each for 10s: a neutral female face (trial 1), followed by a threat face of the same female (trial 2), and a neutral male face (trial 3), followed by a threat face of the same male (trial 4). Prior to each trial the apparatus had a purple cover over the window, and the experimenter would remove the cover so the monkey could see the photo to start each trial. (b) The gaze following task measured co-orienting responses with a demonstrator. Monkeys completed up to four trials where the experimenter captured the monkey's attention and then looked directly upward for 10s; we measured whether the monkey also looked up during this period.

The goal of this study was to assess individual variation in responses, and looking time tasks often produce habituation-related declined in looking over trials, so following prior work we presented the stimuli in a fixed order (Rosati, Arre, et al., 2018). The logic of this experimental design is that, while this fixed order might not be appropriate to assess group-level differentiation of the emotional conditions, it is a more sensitive design to detect relative variation in responses across individuals—without additional variation being introduced by the fact that different monkeys viewed the stimuli in different sequences. As in prior work, monkeys could complete up to four trials, and needed to complete at least the first two trials (involving female photographs) to be included in any

This article is protected by copyright. All rights reserved.

Author Manuscript

analyses. Some monkeys completed only these first trials and are therefore not included in the analyses focusing on responses to male photos. As our prior study showed that monkeys often had quite distinct responses to the female photographs (presented in trials 1-2) and the male photographs (presented in trials 3-4), many of our analyses examined these trials separately.

In test sessions, experimenter one presented the photo stimuli to the monkey using a portable apparatus, while experimenter two filmed the monkey's responses. Specifically, conspecific photos were presented using a white posterboard box (34 cm high, 16 cm deep, 54 cm wide) with a front-facing window to display the photos (23 cm wide, 30 cm tall). This window was initially covered by a purple flap attached with velcro to the main box. When initiating a trial, E1 would rapidly remove the purple cover such that the monkey could observe a photo (see Figure 1a for a diagram of the setup; Supplementary Figure S1 for photographs of the apparatus and example conspecific stimuli; and see Supplementary Video S1 for example experimental demonstrations and monkey looking responses in the task). Between trials, E1 replaced the flap and exchanged the photos out of the monkey's view, and then attracted the monkey's attention to start the next trial as soon as feasible.

On each trial, experimenter one first attracted the monkey's attention to the apparatus by tapping it and calling to the monkey. Once the monkey oriented towards the window, she removed the cover flap to reveal the photo underneath, while simultaneously saying "now;" this auditory cue was used to mark the start of the trial on the video. During each 10s trial, experimenter one held still after initiating the trial, and gazed downwards so her eyes were occluded by a hat. During this time, experimenter two recorded the monkey's face while timing each 10s trial on the camera (see Supplementary Video S1 for an example

This article is protected by copyright. All rights reserved.

Author Manuscript

monkey's response). After 10s were up, experimenter two called "stop" to mark the end of the trial on the video recording.

We presented each subject with a series of photos of conspecifics from the Cayo Santiago population showing another monkey's face and upper shoulders. As in prior work, there were two possible photo sets, each comprising an adult female and an adult male (the age of the male and female individuals approximately matched across the two sets). These two sets allowed us to (mostly) match subject monkeys with a set of unfamiliar conspecifics in the photosets. To do so, we avoided testing subjects with photosets where the subject was living in the photo monkey's natal or current group. The female photos and one of the male photos were the same as in prior work (Rosati, Arre, et al., 2018), but we changed one male set to optimize use of outgroup photos in the current sample. In some cases, however, it was not possible to match a subject with a photoset that did not share either their birth group or current group (or the identity of the monkey was only confirmed after the test was complete). In these cases, we aimed to present photos from individuals who were not current group members, and further controlled for whether outgroup photos were presented in statistical models.

2.5 Gaze following task procedure

This task aimed to characterize individual variation in gaze sensitivity, using procedures that were largely identical to one previously implemented with this population (Rosati et al., 2016). Experimenter one, the demonstrator, first attracted the monkey's attention to her face (by calling to the monkey or snapping her fingers). Once the monkey was judged to be looking, she then looked directly up while simultaneously saying 'now' (see Figure 1b for a diagram of the setup; and Supplementary Video S2 for example

This article is protected by copyright. All rights reserved.

Author Manuscript

experimental demonstrations and monkey responses in the task). As in the socioemotional responses task, this auditory cue was used to mark the start of the trial on the video. Experimenter two, the camera person, stood next to experimenter one and filmed the monkey's face. After 10s trial had concluded, experimenter two said 'stop' to mark the end of the trial on the video. We presented monkeys with up to four identical trials. Across the four trials, experimenter one paused for at least 30s between trials; during this time, she stood looking away from the monkey. This change from prior work at Cayo Santiago without such an inter-trial delay (e.g., Rosati et al., 2016) was aimed at reducing habituation across the four trials, a phenomenon where animals who have flexible control over their responses tend to reduce responding to repeated looks (Rosati et al., 2016; Tomasello et al., 2001). Monkeys had to complete at least the first two trials to be included in the final dataset.

We used a human demonstrator in this task to ensure the demonstrator's behavior was tightly controlled and consistent. As in prior similar work, there was not a specific target of the experimenter's gaze in this context, so we tested monkeys sitting in the vicinity of a tree such that the human's actions were consistent with the possibility of a target of attention being present, as in prior work using similar gaze following methods (Bettle & Rosati, 2019; Rosati et al., 2016). We therefore did not initiate tests when we observed that another monkey was actually present in a tree above the subject, to avoid any possible visual and auditory confounds.

2.6 Exclusions

In this free-ranging context, monkeys were necessarily tested when other monkeys were also present in the vicinity. As such, some monkeys were approached for testing but

This article is protected by copyright. All rights reserved.

Author Manuscript

would fail to produce a scoreable response, typically because they walked away from the testing area before completing both trials, because they were displaced by other monkeys before completing the session, or because the video was not possible to score (such as due to backlighting, or monkeys moving out of the shot). These sessions did not have scorable responses because the monkeys did not complete the number of trials that was necessary for the study. If the monkey left the testing area or was interfered with while the session was ongoing, experimenter one would judge whether or not it was possible to continue to test the monkey on the subsequent trials based on their spatial position, attention to the apparatus, and social context, following general procedures for field experimental studies at Cayo Santiago (see also Huang, Chang, Santos, & Rosati, 2024). For example, in many cases monkeys may have moved to a location where it was not possible to place the apparatus in an appropriate configuration to complete additional trials in the socioemotional task, or the monkey may have moved away from the tree in the gaze task. Note that E1 could make this decision blind to the monkey's specific looking responses in the prior trials, because she was not looking at the monkey the timed 10s trial in either task.

Numbers of exclusions (n = 17 and n = 10 monkeys in the socioemotional responses task year one and two, respectively; n = 6 and n = 5 monkeys in the gaze following task year one and two, respectively) were generally similar to or less than exclusion rates in comparable prior studies in this free ranging population (Bettle & Rosati, 2019, 2021; Hughes & Santos, 2012; Marticorena, Ruiz, Mukerji, Goddu, & Santos, 2011a; Martin & Santos, 2014; Rosati et al., 2016; Rosati, Arre, et al., 2018). In addition, in some cases target animals were approached for testing more than once, because we only confirmed

This article is protected by copyright. All rights reserved.

Author Manuscript

their ID after the test was complete. In this situation, we only included and analyzed their first successful session in the final dataset.

2.7 Coding

Two coders who were both blind to trial type and condition independently scored all trials from the final set of subjects, resulting in both a primary code and a reliability code. Coders were blind to condition because each individual trial was clipped from longer video sessions, randomized across all trials, and renamed with a random number (e.g., clips 1, 2, 3, etc.), and different trials from the same monkey were interspersed across the full set of clips with no identifying cues. A given trial clip started 1-2 seconds before the experimenter one said "now" while she was initially attracting the monkey's attention, which allows coders to better judge where the monkey is looking at the start of the trial as a reference, and ended after experimenter two said "stop" (see Supplementary Video S1 for example clips showing monkey looking responses in the socioemotional responses task, and Supplementary Video S2 for example clips showing monkey responses in the gaze following task).

The coders examined these video clips frame-by-frame using the program Filmora to assess the monkeys' responses following typical methods used in prior studies in this population. For looking time responses, we followed the same coding procedures described in prior looking time studies with the Cayo Santiago population (Bettle & Rosati, 2021; De Petrillo & Rosati, 2019; Huang et al., 2024; Hughes & Santos, 2012; Marticorena, Ruiz, Mukerji, Goddu, & Santos, 2011b; Martin & Santos, 2014). Here, the monkey's initial looking direction at the start of the trial (when E1 said "now") served as the reference to code their looking towards the apparatus for the subsequent 10s. Clips were always coded

This article is protected by copyright. All rights reserved.

Author Manuscript

for 10s from when experimenter one initiated the trial to equate total trial duration across monkeys. For the gaze following task, we followed the coding procedures of prior studies using this kind of setup with the Cayo Santiago population (Bettle & Rosati, 2019; Bettle & Rosati, 2022; Rosati et al., 2016; Rosati & Santos, 2017). This coding method was similar to those for looking time data, but here the initial looking direction was to the demonstrator and we rather judged whether the monkey ever looked upwards (using either their eyes or entire head) as our primary measure.

All coders had previous experience coding monkey data by completing a relevant coding training set to be able to reliably code looking time and/or gaze following tasks in this population. AD was the primary coder for all tasks (in both years); there were different secondary reliability coders for each task in each year (this was different people due to coder availability over time). We calculated reliability for our primary measures in each task in each year to check consistency of the codes. In the socioemotional biases task, coders had high correlation in looking times for both year one (Pearson's r = 0.93), and year two (Pearson's r = 0.97). For the gaze following task, the coders also had high reliability for whether the monkey ever looked up on each trial (Year one Cohen's κ = 0.98; Year two Cohen's κ = 0.95), as well as the total time they spent looking up in the trial (Year one Pearson's r = 0.95; Year two Pearson's r = 0.96).

2.8 Statistical analyses

Our first goal was to assess patterns of variation in social cognition by age and sex. To address this, we analyzed individual variation in performance in each of the cognitive tasks using multi-level models, as in our preregistered analysis plan (https://aspredicted.org/48K_8XP; additional aspects of the pre-registered analysis plan for

This article is protected by copyright. All rights reserved.

Author Manuscript

the entire project focus on associations between cognitive health and social behavior metrics, which are not the focus of the current paper). We used the statistical software program R (R. Core Team, 2022), and implemented mixed models in the *lme4* package (Bates, 2010; Bates, Mächler, Bolker, & Walker, 2015). We report Wald's confidence intervals in parameter reporting using the function *confint* from that package. Across tasks, we used likelihood ratio tests (Bolker et al., 2008) to first compare base versus full model fit, and then tested the sequential inclusion of specific predictor variables of interest. We conducted *post hoc* Tukey comparisons of model factors using the *emmeans* package (Lenth, 2018). We checked that models met assumptions (homogeneity and normality of residuals) using the *DHARMa* package in R (Hartig, 2022).

Depending on the structure of the cognitive data, we either used the *glmer* function to implement generalized linear mixed-effects models (GLMMs), or the *lmer* function from lme4 to implement linear mixed-effects models (LMMs). In the socioemotional responses task, initial checks showed that looking time data was fairly right skewed, so we used GLMMs implemented with a log-link gamma distribution as is recommended for such data (Lo & Andrews, 2015). Such a modeling approach is similar to log-transforming the data but rather implements generalize models that account for the right-skew. For analyses of difference scores in this task, we used LMMs because the distribution was fairly normal. In the gaze task, the primary measure was a binary outcome, so we implemented GLMMs with a binomial distribution. In some secondary analyses we used duration of looking as a proxy for co-orienting responses; we used gamma GLMMs to model this data as it was also right-skewed. Some models examining looking time responses still did not meet all model assumptions even when accounting for right-skew in this way, but we found largely

This article is protected by copyright. All rights reserved.

Author Manuscript

similar results implementing these tests with a different structure, (e.g., using LMMs) suggesting these patterns were generally robust to the model structure choices.

For our second goal of assessing performance over time and across tasks, we completed several additional analyses. First, we examined individuals who completed the same task at both time points, using the *rptR* package in R to quantify long-term repeatability in performance. For a given task, we estimated repeatability after accounting for age and sex for each trial separately (note that these trials measured distinct aspects of responses, and that inclusion of multiple trials from the same task would rather measure short-term repeatability within the same testing session). Here, we used LMMs for the socioemotional responses as this was the most appropriate option implemented in this package, and binomial GLMMs for the gaze task. Second, we further used mixed models to test if the first year's cognitive response was a significant predictor of the second year's responses in that task, using the same model structure as for the main task assessments. This allowed us to address stability in responses over time using a converging approach where multiple trials could be assessed holistically.

Then, we examined within-individual (longitudinal) age-related changes in cognitive performance. To do so, we estimated repeated measures correlations implemented with the R package *rmcorr*, which tests for a common within-individual association for paired measures taken at two or more time points for multiple subjects (Bakdash & Marusich, 2017). Repeated measures correlation is a method that is close to a null multilevel model of varying intercept and a common slope for each individual, but specifically tests for the association shared among individuals (i.e., common regression slope), after adjusting for inter-individual variability (Bakdash & Marusich, 2017). We

This article is protected by copyright. All rights reserved.

Author Manuscript

used this to examine within-subject (longitudinal) change in looking times with age, as our primary mixed model analyses incorporate both longitudinal and cross-sectional age-related changes. For each trial, we first used repeated measures correlations to isolate the common intra-individual variance in looking time across age reflecting longitudinal changes. We further constructed a LMMs with random intercepts for individual identity to estimate age effects in the same sample. This allowed us to report quantitative estimates of both overall age-related change (incorporating both longitudinal within-subjects and cross-sectional between-subjects age shifts),to compare with the common longitudinal change estimated from the repeated measures correlations.

Finally, for our third goal we assessed if performance across the two different social tasks were related to evaluate if these tasks recruited overlapping versus distinct cognitive processes. For this, we examined the subset of individuals who completed both tasks in the same year and used performance in one task to predict performance in the other, Since animals may have only completed the first two trials in each task, when that task was used as a predictor variable for the other task, we averaged just the first two trials that all animals had completed to make this predictor equitable across individuals. We then used GLMMs to predict responding in one task based on performance in the other, using the same approach as the primary modeling of tasks responses but then adding performance in the other task to test if it further improved model fit.

We analyzed these data with age in years as a continuous predictor, but in some figures we split individuals into age cohorts: younger monkeys up to 15 years; and older monkeys over 15 years, as monkeys in this population have a median lifespan of 15 years and rarely exceed 25 years (Hoffman, Higham, Mas-Rivera, Ayala, & Mastripieri, 2010).

This article is protected by copyright. All rights reserved.

Author Manuscript

While prior work also included infant and/or juvenile cohorts comprising individuals under the age of five years (Rosati et al., 2016; Rosati, Arre, et al., 2018), we tested very few monkeys between the ages of 4-5 years in the current study and included those individuals in the younger adult cohort for visualization purposes.

2.9 Data availability statement

Data will be made publicly available in Dryad Data Repository upon publication, and is currently available for review at https://datadryad.org/stash/share/fFMqP5ZG-h0upPgoZ2XfbD5Ttxy4aXJA5emRrQudMvA.

3. Results

3.1 Individual variation in the socioemotional responses task

We first examined duration of looking to the photo stimuli across the two years of data. There were 181 total monkeys in the sample (all of them completed at least the first two trials presenting female photos in at least one year of testing; see Table 1 for breakdown of participation across years). Overall, monkeys looked an average of M = 2.35 ± SE = 0.13s at the female neutral photo; 1.90 ± 0.11s at the female threat photo; 1.88 ± 0.14s at the male neutral photo; and 2.19 ± 0.15s at the male threat photo (see Figure 2a for breakdown by subject's sex and age cohort). In an initial analysis of the full dataset, we implemented GLMM models with gamma-log link distribution. The base model accounted for *photo set* (which of the two sets of stimuli was presented), *outgroup photos* (whether monkeys saw unfamiliar conspecifics as the female and male stimuli), and subject's *identity* as a random intercept. We then compared this to a full model including (1) the interaction between *face stimuli (*female versus male photos) X *expression* (neutral versus threat photos), the key index of the different kinds of stimuli presented in the task; (2)

This article is protected by copyright. All rights reserved.

Author Manuscript

subject's *age*; and (3) subject's *sex* to test if these primary variables of interest predicted response. Indeed, this improved fit [$\chi2 = 68.79$, df = 5, $p < 0.0001$]. We then checked the sequential addition of these variables. There was significant improvement with the inclusion of the *face X expression* interaction, the key index of the different stimuli presented, compared to the base model [$\chi2 = 24.96$, df = 3, $p < 0.0001$]. Post hoc tests indicated that monkeys generally looked longer at the female photos than the male photos ($p < 0.0001$), which is not surprising since they were presented first in the sequence., but also that while monkeys looked significantly less at the female threat photo compared to the female neutral photo ($p< 0.0001$), they showed no such decline for the male threat photo compared to the male neutral photo ($p = 0.16$, n.s.). The inclusion of age also improved fit [$\chi2 = 43.83$, df = 1, $p < 0.0001$]: monkeys looked less at photographs as they aged. Further inclusion of subject's *sex* did not [$\chi2 = 0.004$, df = 1, $p = 0.95$, n.s.], indicating no differences between the responses of male and female monkeys. Overall, this initial analysis indicated that monkeys show age-related shifts in looking during adulthood, but also qualitatively different reactions to the male versus female emotional stimuli

Given that these different patterns of responding to the female versus male stimuli mirrored those seen in prior work, and that distinct subsets of monkeys completed either just the first two female stimuli trials versus the male stimuli trials as well, we then examined more nuanced aspects of these age-related shifts in separate analyses of responses to the female and male stimuli, using the same analytic approach from prior work with this task (Rosati, Arre, et al., 2018). To examine responses to the female stimuli, we implemented a base GLMM model accounting for *photo set* (which of the two sets of stimuli was presented), *outgroup photos* (whether monkeys saw unfamiliar conspecifics as

This article is protected by copyright. All rights reserved.

Author Manuscript

the female stimuli), and subject's *identity* as a random intercept. We then compared this to a full model including (1) photo e*xpression* (neutral versus threat); (2) subject's *age;* (3) subject's *sex*; and (4) an *expression X age* interaction, the main test of whether socioemotional biases change across the lifespan. The full model showed improved model fit compared to the base model [$\chi2 = 67.51$, df = 4, $p < 0.0001$]. We then checked the sequential addition of these variables. First, there was significant improvement with the inclusion of the *expression* compared to the base model [$\chi2 = 15.08$, df = 1, $p = 0.001$]: monkeys viewed the female neutral photo longer than the threat photo. Adding in *age* further improved fit [$\chi2 = 47.15$, df = 1, $p < 0.0001$]: overall, monkeys looked less at photos with increasing age. In contrast, there was no improvement with inclusion of *sex* [$\chi2 = 0.83$, df = 1, $p = 0.36$, n.s.], indicating no differences in responses between males and females. Finally, the *expression X age* interaction further improved fit [$\chi2 = 4.45$, df = 1, $p = 0.034$; see supplement for model parameters]. Post-hoc tests revealed a significant difference in age slopes for responses to the neural versus threat photo [$p = 0.034$], such that older monkeys an attenuated a decline in looking at the threat compared to the neutral photo. Additional checks confirmed that these patterns held for monkeys tested with both photo sets, with no differences by photo set that was presented.

We used the same approach to examine looking to the male photo stimuli. As for female photos, the inclusion of all experimental variables improved fit in the full model compared to the base model [$\chi2 = 18.29$, df = 4, $p = 0.001$]. We then checked the sequential addition of these variables. Here, inclusion of *expression* did not improve model fit compared to the base model, unlike with the female photos [$\chi2 = 1.52$, df = 1, $p = 0.22$, n.s.]. Further inclusion of *age* did improve fit compared to the base model [$\chi2 = 16.96$, df

This article is protected by copyright. All rights reserved.

Author Manuscript

= 2, p = 0.0002]: overall, monkeys looked less at photos with increasing age. There was again no improvement with inclusion of *sex* [$\chi 2 = 0.54$, df = 1, p = 0.46, n.s.], indicating no differences between males and females. We finally added the *expression X age* interaction, which unlike with the female photos did not improve fit [$\chi 2 = 0.78$, df = 1, p = 0.38, n.s.; see supplement for model parameters]. We also checked if these results depended on the photoset the monkeys viewed, and here found that this did impact results: inclusion of an interaction between expression and photoset improved model fit [$\chi 2 = 22.63$, df = 1, $p < 0.0001$]; post-hoc tests showed that whereas monkeys looked longer at the threat than the neutral expression if they viewed photoset 1 (p = 0.0001) aligning with prior work, they looked longer at the neutral photo if they were presented with photoset 2 (p = 0.0022). Thus, in this dataset there were differences in responses to male photos depending on the identity of the male photo presented to the subject monkeys.

We next calculated a *difference score* (Threat looking time – Neutral looking time) in order to examine overall age-related changes to emotional stimuli and compare responses to female versus male photos, following prior work (Rosati, Arre, et al., 2018). This sort of difference score is commonly used in looking time research with human infants (Kominsky, Li, & Carey, 2022; McCrink & Wynn, 2004; Oakes, 2017; Olineck & Poulin-Dubois, 2009; Spelke, Breinlinger, Macomber, & Jacobson, 1992). Here, this score indexed how monkeys allocated their looking at the threat photo while accounting for overall looking to a neutral expression of the matched conspecific, such that more positive scores indicate greater relative interest in the negative threat photo. On average, we found that the difference scores were negative for the female photos (M = -0.42 ± SE = 0.11), indicating longer relative looking to the neutral photo in that set, and positive for male

This article is protected by copyright. All rights reserved.

Author Manuscript

photos (0.30 ± 17) indicating longer relative looking to the threat photo. To analyze this, we constructed a base LMM accounting for *photo set*, *outgroup* photos, and subject *identity* as a random factor; the full model including *photo type* (male or female photos), subject's *age*, subject's *sex*, and the interaction between *age X photo type* improved model fit [$\chi 2$ = 19.73, df = 4, p = 0.0006]. We then looked at each predictor sequentially. Inclusion of *photo type* (male or female photos) improved fit [$\chi 2$ = 11.82, df = 1, p = 0.0006]; monkeys indeed had more positive scores to the male photos. While the inclusion of age only trended to improve fit [$\chi 2$ = 2.90, df = 1, p = 0.09], the further inclusion of both *age* and *sex* improved fit compared to the model without individual subject characteristics [$\chi 2$ = 7.01, df = 2, p = 0.03]: male subjects tended to have lower scores than females, and scores generally become more positive (indicating greater relative attention to the threat photograph) with increasing age. Finally, we examined if there was any difference in this trajectory for male versus female photos by including the *age X photo sex* interaction; this did not further improve fit [$\chi 2$ = 0.90, df = 1, p = 0.34, n.s.; see Figure 2b and supplement for model parameters] indicating similar age-related increases in attention to threat for both male and female stimuli as measured by the difference score.

3.2 Stability and longitudinal changes in socioemotional responses

We then examined patterns of cognition across the subset of animals that completed the socioemotional responses task in both years of data collection. First, we tested for repeatability of individual responses across repeated testing, while also controlling for age and sex. We examined repeatability for each trial separately, as these trials involved different kinds of stimuli (e.g., neutral versus threat, and female versus male photos), and we also were interested in capturing long-term repeatability in contrast to between trial

This article is protected by copyright. All rights reserved.

Author Manuscript

repeatability rather than repeatability over successive trials. We found significant long-term repeatability for responses to female neutral photos, female threat photos, and male threat photos over the one-year period (see Table 2 for statistical reporting). Notably, these estimates were stronger for responses to the threat photos for both kinds of stimuli.

We further used GLMMs to test how performance in year one predicted performance in year two in a more holistic fashion, examining the multiple test trials in tandem. Here, we used the subset of data where a given monkey completed the same trials in both years. We first constructed a base gamma GLMM model of year two looking times that included the same variables used the prior analyses (e.g., accounting for subject's *identity*, *age* in year two, *sex*, the *photoset*, *outgroup*, and the interaction of *photo type* and *expression*). In the second model, we then further added *year one looking time* (e.g., for the same stimulus) to test if individual looking time performance in year one predicted that of year one. This improved fit [$\chi 2 = 8.806$, df = 1, $p = 0.003$; see Figure 2c and supplement for model parameters]: above and beyond our original predictors, year one predicted year two performance. This indicates that year one performance was similarly predictive for looking in year two for both neutral and negative photos. Overall, this provides converging evidence that individual variation in cognitive responses were stable across testing.

This article is protected by copyright. All rights reserved.

Author Manuscript

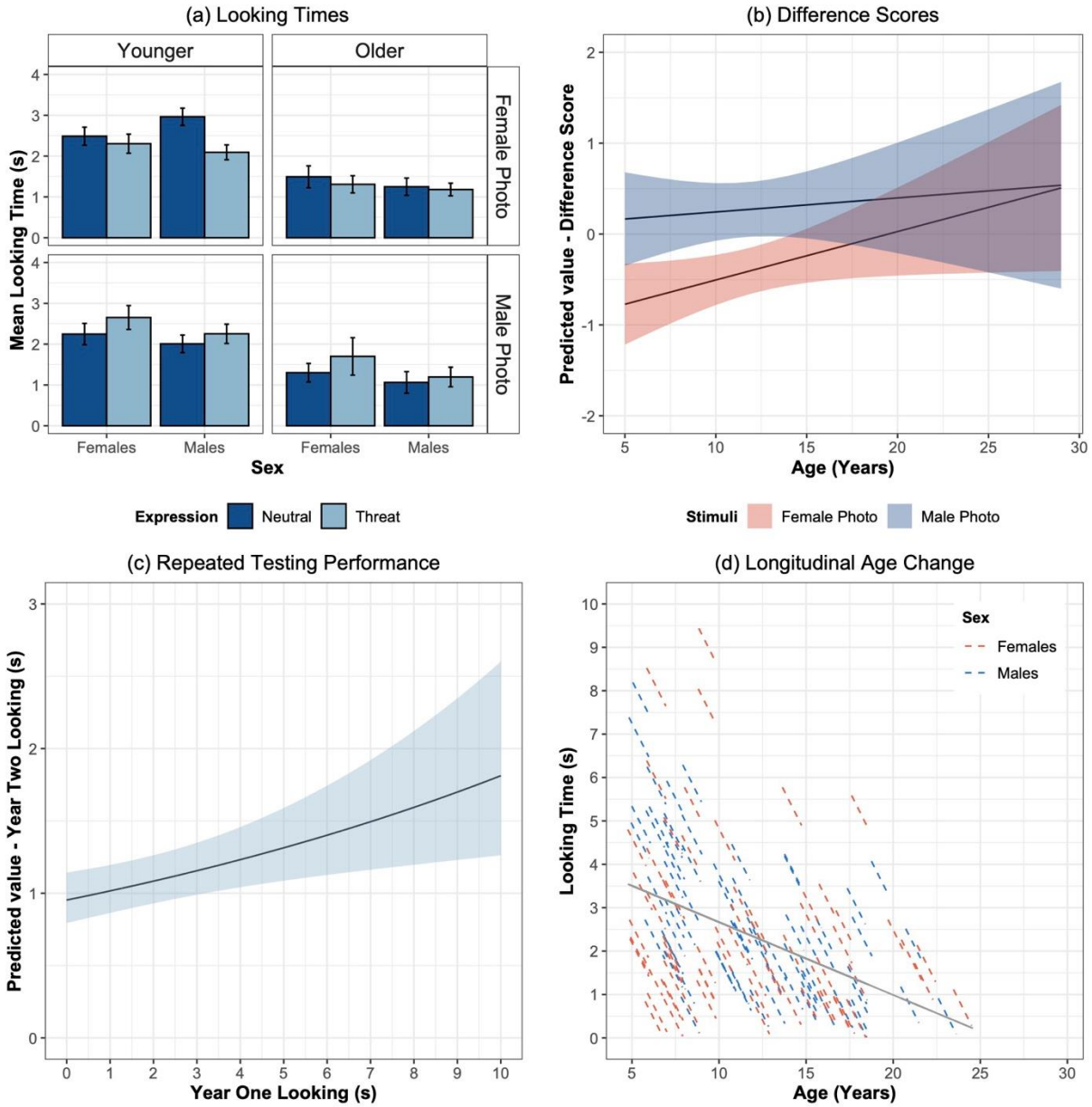


Figure 2: Individual variation in the socioemotional responses task. (a) Mean looking time to conspecific neutral and threat photographs. Graph shows mean duration of looking time towards neutral versus threat stimuli for male and female photographs, broken down by subjects' age cohort and sex. Error bars indicate standard error. (b) Predicted changes in difference scores (negativity bias) with age. Each subject was assigned a difference score indexing their negativity bias (threat looking time–neutral looking time) for female and male photographs, which increase with age. Estimated values are derived from linear mixed models also accounting for subjects' identity, sex, age, and experimental variables. Ribbons indicate 95% confidence intervals. (c) Predicted performance over repeated testing. We estimated how performance in year two of testing was related to year one performance for individuals who completed both years of testing. Estimated

This article is protected by copyright. All rights reserved.

Author Manuscript

values are derived from linear mixed models also accounting for subjects' identity, sex, age, and experimental variables. Ribbons indicate 95% confidence intervals. (d) Longitudinal changes in looking responses. We used repeated measures correlations to isolate within-individual associations in looking time. The panel depicts responses to the first trial (female neutral photos). The solid gray line indicates the overall linear relationship between age and looking time in the dataset incorporating both within- and between-individual age effects (estimated from a mixed model), whereas the dashed lines indicate the common within-individual age-related change estimated from repeated measures correlation.

Finally, we used this same subset of monkeys to examine longitudinal (within-individual) age changes in responses to emotional stimuli. Specifically, we used repeated measures correlations to assess common within-individual associations (Bakdash & Marusich, 2017). For example, repeated measures correlations revealed a within-individual steeper decline in looking time with age [age slope = -0.80, r = -0.30, p = 0.0001] compared to equivalent LMMs estimates [estimate = -0.17, SE = 0.03, t= -6.44, p < 0.0001; see Figure 2d]. We found similar patterns of exacerbated within-individual age declines for the *female threat photo* [repeated measures correlation: age slope = -0.89, r = -0.39, p < 0.0001; LMM age estimate: -0.11, SE = 0.03, t= -4.46, p < 0.0001], the *male neutral photo* [repeated measures correlation: age slope = -0.62, r = -0.25, p = 0.024; LMM age estimate: -0.15, SE = 0.04, t= -4.20, p < 0.0001], and the *male threat photo*, although this final trial did not show a significant within-individual age effect [repeated measures correlation: age slope = -0.46, r = -0.18, p =0.10, n.s.; LMM age estimate: -0.12, SE = 0.04, t= -2.65, p = 0.01]. Overall, this indicates a stronger negative association between age and looking time at the intra-individual (longitudinal) level, compared to the more moderate mean negative relationship between age and looking time across monkeys estimated from mixed models incorporating both longitudinal and cross-sectional age estimates.

3.3 Individual variation in gaze following task

This article is protected by copyright. All rights reserved.

Author Manuscript

We first examined the proportion of monkeys following gaze across the two years of data. Overall, M = 38.6 ± SE = 2.9% of monkeys looked up on the first trial, 31.8 ± 2.8% of monkeys looked up on the second trial, 23.4 ± 2.8% of monkeys looked up on the third trial, and 22.2 ± 3.1% of monkeys looked up on the last trial (see Figure 3a for breakdown by age and sex). Following prior work (Rosati et al., 2016), we analyzed propensity to look across trials as our primary measure. We implemented a base model accounting for subject's *identity* as a random factor, and then in the full model included (1) *trial number* (as an ordinal factor, to test for habituation in responses over trials), (2) subject's *age*, (3) subject's *sex*; and (4) the interaction between *age* X *trial number*. This improved model fit compared to the base model [$\chi 2 = 23.13$, df = 4, $p < 0.0001$]. We then examined these factors sequentially. Inclusion of trial number improved fit compared to the base model [$\chi 2 = 23.42$, df = 1, $p < 0.0001$], revealing a significant effect of trial number: monkeys looked less across successive trials, indicating that they habituated to the repeated gazes of the actor. However, the subsequent additions of *age, sex,* and an *age X trial* interaction did not further improve fit [$p > 0.5$, n.s. in all cases; see Figure 3b and supplement for model parameters], indicating little variation in gaze following across individuals according to their demographic characteristics.

We next examined patterns of looking when the monkey did produce a look (i.e., only those trials where the monkey looked up). Overall, monkeys spent an average of 1.58 ± 0.09s looking up when they produced a look, with similar durations of looking across trials. We confirmed this by analyzing duration of looking. Similar to the primary analyses of propensity to look, we implemented a base model accounting for subject's *identity* as a random factor, and then in the full model added *trial number, sex,* and *age* to subsequent

This article is protected by copyright. All rights reserved.

Author Manuscript

models. There was no improvement in model fit for the full model [$\chi 2 = 2.02$, df = 3, $p = 0.56$, n.s.; see supplement for model parameters], and the individual inclusion of these factors also not improve model fit. Overall, this indicates that if monkeys did look up, the temporal characteristics of their looking responses were similar across trials, sexes, and cohorts.

3.4 Stability and longitudinal changes in gaze following responses

We next examined patterns of cognition across the subset of animals that completed the gaze following task in both years of data collection. We used the same approach to quantifying long-term repeatibility as for the socioemotional responses task, here using a binomial models, and again accounting for age and sex to model responses in each separate trial (note that these characterstics were removed in the analyses of trial 4 so that models could converge, given low sample size). This allowed us to test if individuals who looked up in the first year of testing also did so in the second year (rather than repeatibility over successive trials in the same year). In fact, we found significant repeatability for the first trial, but not the last three trials (although there was a trend for trial 4; see Table 3).

This article is protected by copyright. All rights reserved.

Author Manuscript

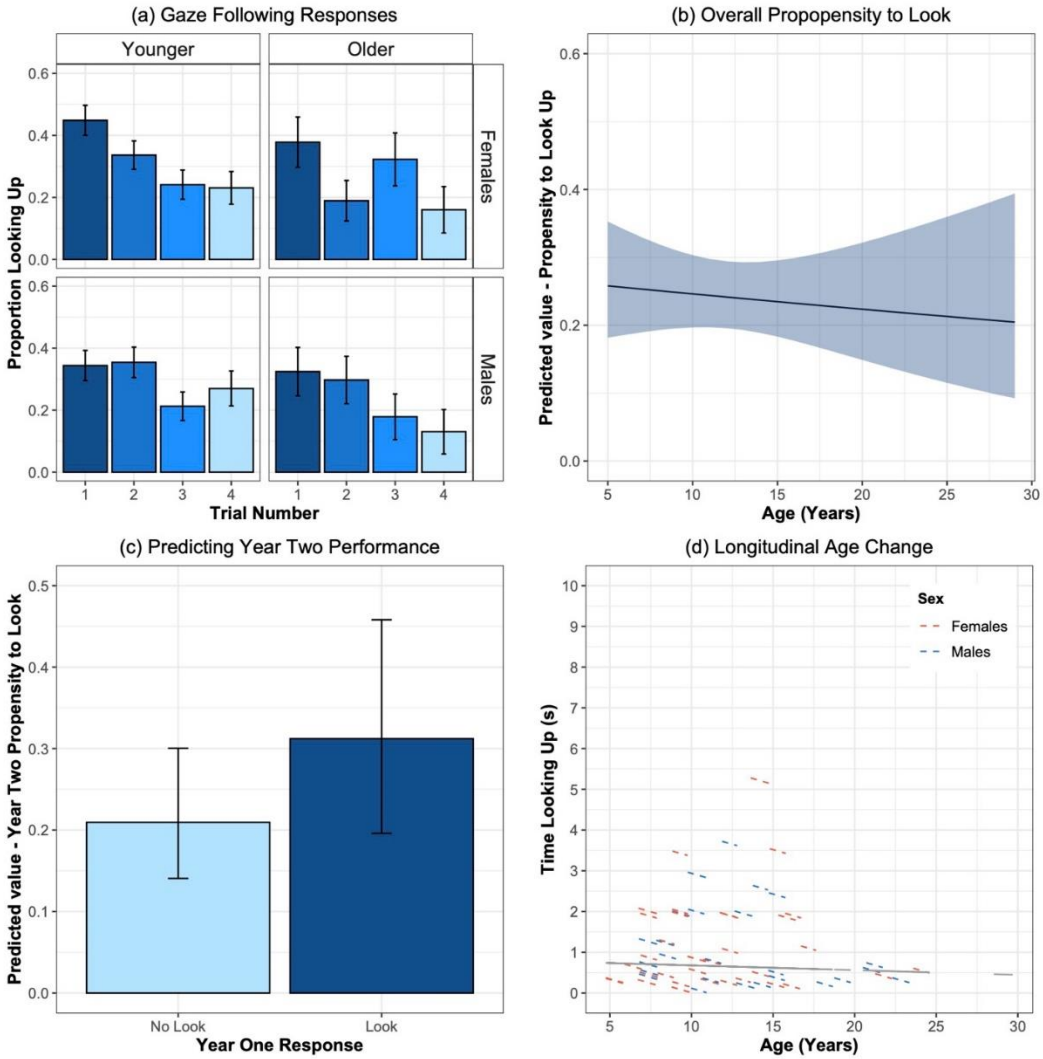


Figure 3: Individual variation in the gaze following task. (a) Propensity to follow gaze. Graph shows mean proportion of individuals looking up across the four trials, broken down by subjects' age and sex. Error bars indicate standard error. (b) Predicted propensity to follow gaze did not change with age. Estimated values are derived from GLMMs of propensity to look up also accounting for subjects' identity sex, age, and trial number. Ribbons indicate 95% confidence intervals. (c) Predicted performance over repeated testing. We tested if performance in year two was related to year one performance. There was no relationship between performance in year one and likelihood of looking up in year two for individuals who completed both years of testing. Estimated values are derived from GLMMS of propensity to look up accounting for subjects' identity, sex, age, and trial number. Ribbons indicate 95% confidence intervals. (d) Longitudinal changes in duration of looking responses. We used repeated measures correlations to isolate within-individual associations in time looking up. The panel depicts responses to the first trial. The solid gray line indicates the (lack of) overall linear relationship between age and time spent looking up,

This article is protected by copyright. All rights reserved.

Author Manuscript

whereas the dashed lines indicate the (lack of) within-individual changes with age.

We further used GLMMs to see how performance in year one predicted year two in this subset of the monkeys. Specifically, we created a base model of year one gaze following responses (as a binary variable) accounting for the additional factors from our prior models: subject's *identity*, *age* in year two, *sex*, and *trial number*. We then added *year one response* (look or no look) for that trial in a second (full) model to test if individual performance was consistent over time. This did not improve fit [$\chi 2 = 2.26$, df = 1, p = 0.13, n.s.; see Figure 3c and see supplement for model parameters]. Overall, this aligns with the results from the repeatability analyses, indicating that an individual's gaze following responses were not particularly stable across repeated testing a year apart.

Finally, we used this same subset of monkeys to examine longitudinal (within-individual) age changes in gaze following, using repeated measures correlations as in the socioemotional responses task. As repeated measure correlations can only use continuous data, for this analysis we used duration of looking up as a proxy for gaze following responses. Unlike in the socioemotional responses task, there was generally no significant effect of age using either approach. Repeated measures correlations were not different from 0 [e.g., for trial 1: repeated measures correlation: age slope = -0.12, r = -0.07, p = 0.50, n.s.; see Figure 3d], nor were LMMs estimates [for trial 1: age estimate = -0.012, SE = 0.020, t= -0.608, p =0.54, n.s.]. This aligns with the results from the primary analyses, which also indicated no age effects, and further shows that observed declines were not detectable within-individuals either.

3.5 Relationship between sociocognitive skills

This article is protected by copyright. All rights reserved.

Author Manuscript

We finally further used mixed models to test whether and how performance in one task was related to performance in the other in the same year of testing. First, we used gaze following to predict looking to conspecific faces in the socioemotional responses task, focused on the 151 monkeys who completed both tasks at least at one timepoint. We created a base gamma-log link GLMM of looking times in the socioemotional task, mirroring the initial models that accounted for subject's *identity*, *age* at the time of the emotions task, *sex*, *photoset*, *outgroup photos*, and the interaction of *photo type* (female or male photos) and *expression* (neutral or threat). We then added *average gaze following responses* in a second (full) model, which did not improve model fit [$\chi 2 = 1.31$, df = 1, p =0.25, n.s.; see Figure 4a and supplement for model parameters]. We then examined the reverse relationship by testing if attention to conspecific photos predicted gaze following. Specifically, we created a base binomial GLMM model of gaze following responses accounting for subject's *identity*, *age* at the time of the gaze task, *sex*, and *trial number*. We then added *average looking time* in a second (full) model, which here did improve fit [$\chi 2 = 9.14$, df = 1, p = 0.002; see Figure 4b and supplement for model parameters]. Overall, this shows that individual monkeys who showed more sustained interest in conspecific photographs were more likely to follow the gaze of a human actor, although propensity to engage in gaze following was a less-reliable predictor of responses to the conspecific photos.

This article is protected by copyright. All rights reserved.

Author Manuscript

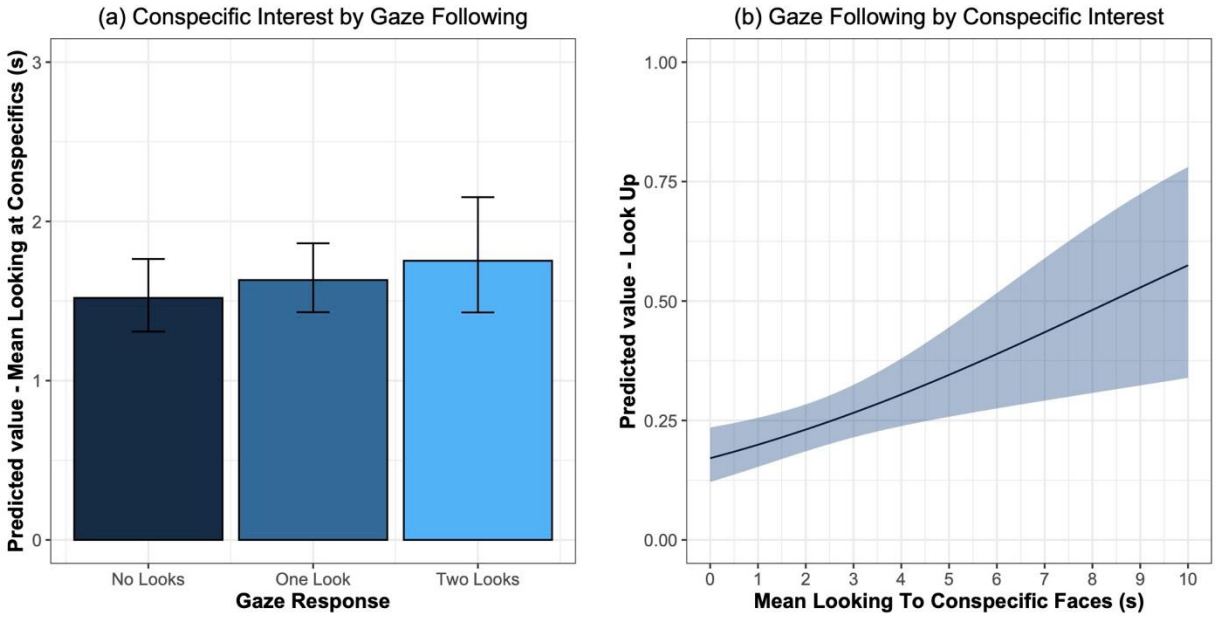


Figure 4: Inter-task relationships. (a) Predicted performance in the socioemotional responses task from the gaze following task. Responses to conspecific photos for individuals who completed both tasks in a given year are depicted, broken down by their responses in the first two trials of the gaze following task (which all individuals completed; they therefore could have gaze followed on up to two trials). Gaze following did not significantly predict interest in the conspecific photographs. Estimated values are derived from linear mixed models also accounting for subjects' identity sex, and age, and experimental variables. Error bars indicate 95% confidence intervals. (b) Predicted performance in gaze following task from the socioemotional responses task. Responses to gaze following demonstrations for individuals who completed both tasks in a given year are depicted, according to their average duration of looking at conspecific photos in the two trials of the socioemotional responses task (which all individuals completed). Individuals who looked longer at conspecific photos were more likely to look up in the gaze following task. Estimated values are derived from GLMMS of propensity to look up also accounting for subjects' identity, sex, age, and trial number. Ribbons indicate 95% confidence intervals.

4. General discussion

We used a longitudinal approach to measure different aspects of social cognition in a large sample of rhesus monkeys across the adult lifespan at two time points. In the socioemotional responses task assessing attention to emotionally-valanced versus matched neutral images, we found that monkeys exhibited general declines with age in social

This article is protected by copyright. All rights reserved.

Author Manuscript

attention but also showed increases in relative attention to negative threat stimuli. These age-related shifts were detectable also in longitudinal analyses. Our analysis showed that individual differences in performance in the task was repeatable and stable over repeated testing. In the gaze following task, we measured monkeys' co-orienting response to an individual who looked up. Here, we did not find major age-related differences either in the main analyses or in longitudinal-only comparisons. While there was some limited evidence for repeatability, performance was not generally stable within individuals over time. Finally, analyses looking at the inter-relationship of responses across these two tasks revealed that performance across these two distinct facets of social cognition was related: monkeys that looked longer at stimuli in the socioemotional responses task were more likely to co-orient in the gaze following task (although gaze following did not similarly predict looking time responses). Overall, these results show that cognitive tasks implemented with free-ranging animals can capture both age-related changes and stable individual differences, but also that this can depend on the task in question.

Our findings of patterns of individual variation in the socioemotional responses task largely accorded with prior work using this task (Rosati, Arre, et al., 2018). Specifically, we found that the macaques generally exhibited declines with age in their attention to the conspecific photographs, but this was attenuated for responses to the negative threat expression in older age resulting in relative increases in attention to the threat photographs, with little variation by sex. This same general pattern was found in the analyses of looking time durations across trials, in comparisons of difference scores indexing an individual monkey's relative interest in the emotional threat stimuli, and finally in the longitudinal comparisons of the same monkey's responses over time. However, the current study did

This article is protected by copyright. All rights reserved.

Author Manuscript

find more mixed responses to the male photographs compared to prior work (Rosati, Arre, et al., 2018), as here responses to these photographs varied across the two possible photo sets. Overall, these results align with prior fundings showing that monkeys exhibit age-related shifts in socioemotional attention that differ from patterns in humans, such that older monkeys show more relative interest in negative threat faces. Thus, while older macaques may retain a general interest in other's social interactions (Almeling et al., 2016), they also show an increasing negativity bias in several aspects of cognition and behavior (Machanda & Rosati, 2020). Notably, prior work with this task included a larger sample of monkeys (n = 337) including a cohort of juvenile individuals, whereas the current work shows that these same types of age-related shifts in emotional biases are detectable in even in a sample comprised only of adults.

In the gaze following task, however, we found a different pattern of individual differences than seen in prior work (Rosati et al., 2016). While adults showed habituation over repeated successive trials as expected, there was no decline in propensity to follow gaze with increasing age, and adult females were not more responsive than adult males to the gaze cues. A lack of age-related decline was found both not only in the cross-sectional analyses comprising all data, but also in the longitudinal comparisons of individuals who were tested at two time points. As prior comparisons detecting sex differences and declines in propensity to follow gaze comprised both a much larger sample size (n = 481) and used a cross-sectional approach with a larger cohort of infants and juveniles (Rosati et al., 2016), it is possible that that these effects may require larger sample size to detect, or the inclusion of younger monkeys may heighten any such age-related changes. However, studies of gaze following in this population similarly found age-related declines in more comparable adult

This article is protected by copyright. All rights reserved.

Author Manuscript

samples (Bettle & Rosati, 2019; Bettle & Rosati, 2022), suggesting that there may also be a true shift in gaze following patterns compared to prior work. Notably, the current data was collected after Hurricane Maria, a major natural disaster in the region, whereas those prior studies were conducted before this event. Accumulating evidence shows Hurrian Maria and other severe hurricanes have impacted some aspects of social behavior, life history outcomes, and biological aging in the Cayo Santiago population (Diaz, Steiner, Tuljapurka, Zuo, & Hernandez-Pacheco, 2023; Morcillo, Steiner, Grayson, Ruiz-Lambides, & Hernandez-Pacheco, 2020; Testard et al., 2021; Watowich et al., 2022). Thus, one possibility is that patterns of co-orienting also shifted in the Cayo Santiago rhesus macaques, such that older adults show higher attention or vigilance to gaze cues than did individuals prior to the hurricane. However, it is unclear why the hurricane might have impacted patterns of responding in the gaze following task but not in the socioemotional responses task, so this issue bears further study.

A novel aspect of the current work was the longitudinal design testing individuals at two time points a year apart, allowing us to examine repeatability and stability of cognitive performance in these tasks over a long timescale. In fact, responses in the socioemotional responses task were stable and repeatability over this period, especially responses to the threat photos. While there is some increasing recent evidence for stability in some cognitive responses in animals (Cauchoix, 2018), most of this work has not focused on nonhuman primate species and examined repeatability over shorter time frames than we did here. Our work therefore provides first demonstration of such repeatability in a primate study using looking time methods over a fairly long (one-year) time frame: even these quick cognitive assessments, taking only a few minutes, can capture at least some

This article is protected by copyright. All rights reserved.

Author Manuscript

stable aspects of individual's cognitive responses. One possibility is that we found relatively high repeatability in the socioemotional responses task because it measures of attention to biologically-central stimuli: conspecific faces and emotional signals. Faces in general, and threats specifically, tend to quickly capture attention (Adolphs, 2008; Dal Monte, Costa, Nobel, Murray, & Averbeck, 2015; Hoffman et al., 2007; Lacreuse, Schatz, Strazzullo, King, & Ready, 2013), and may therefore be less influenced by contextual factors that other psychological processes.

However, we found less evidence for individual stability in co-orienting responses in the gaze-following task: only responses on the first trial were repeatable, and there was no evidence that gaze following in the first year of testing overall was related to year two performance. This differing pattern might be related to the more complex nature of the task—which required not only attending to social agents, but also interpreting their social cues in order to co-orient towards a distant target location in the external environment. Yet it is also possible that this different result is related to the smaller sample size assessed for repeatability on successive trials. Indeed, while some prior work with great apes has shown that individual variation in gaze following is stable over time (Bohn et al., 2023), that study involved more trials per individual than ours. As such, gaze following task may have provided less individual resolution compared to the looking time task. More generally, increasing attention to the issue of stable responses in cognition has revealed several cognitive and behavioral metrics that are not stable over time (Cauchoix, 2018; Mason et al., 2021; Rohrer & Ferkin, 2020; Soha, Peters, Anderson, Searcy, & Nowicki, 2019), highlighting that repeatability in cognitive performance cannot necessarily be assumed.

This article is protected by copyright. All rights reserved.

Author Manuscript

Another major finding from the current study is that some aspects of performance across the two tasks was related: individual monkeys that showed longer looking at conspecific faces in the socioemotional responses task were more likely to follow the gaze of a human in the same year of testing, although we did not find support for the reverse relationship. This indicates that there may be some common underlying psychological trait that is reflected in performance in both tasks. Notably, prior work using a large set of tasks in primates has fairly consistently shown that there is often little commonality in responses to different social cognitive, tasks. For example, work assessing a large cognitive task battery found that chimpanzees exhibit a 'spatial' factor reflecting performance on tasks involving spatial cognition, but no specific factor reflect social cognitive performance across different tasks (Herrmann, Hernández-Lloreda, et al., 2010). Similarly, cross-task relationships for great ape cognitive performance were lower when gaze following was included compared to when only non-social tasks were assessed (Bohn et al., 2023). Some work has found no clear evidence that any cognitive tasks show shared variance in other nonhuman primate species (Völter et al., 2022). One possibility is that the interrelationships in the current tasks are in part capturing interest in and attention to social agents, a common aspect of both social cognitive tasks implemented here and perhaps a more low-level aspect of behavior. In any case, this finding should be investigated further, given that the relationship we found was inconsistent.

One important issue to note is that the two tasks implemented here used quite different experimental measures. Both looking time and gaze following methods are commonly used with free-ranging primates, given that both are appropriate for fast assessments of free-ranging animals (Bettle & Rosati, 2019, 2021; Bettle & Rosati, 2022;

This article is protected by copyright. All rights reserved.

Author Manuscript

De Petrillo & Rosati, 2019; Drayton & Santos, 2015; Drayton & Santos, 2017; Dubuc et al., 2016; Higham et al., 2011; Huang et al., 2024; Hughes & Santos, 2012; Marticorena et al., 2011b; Martin & Santos, 2014; Martin & Santos, 2016; Rosati et al., 2016; Rosati, Arre, et al., 2018; Rosati & Santos, 2017; Teufel et al., 2010; Winters, Dubuc, & Higham, 2015). Yet these tasks also have distinct methodological assumptions that should be carefully considered when interpreting results. For example, the socioemotional responses task used a preferential looking time task to measure monkey's interest in various stimuli, and increasing body of work shows that older monkeys generally look for shorter overall durations at a wide variety of both social and nonsocial stimuli (Bettle & Rosati, 2021; De Petrillo & Rosati, 2019; Huang et al., 2024). This highlights the importance of examining patterns of *relative* looking across conditions or stimuli, as we did here. In contrast, in the gaze following task the main measure is whether animals co-orient with a demonstrator. Prior work has also shown that such responses do reflect co-orienting in this population—in that they look in the same direction as the actor and show low baseline rates of looking up in control conditions (Bettle & Rosati, 2019; Bettle & Rosati, 2022; Drayton & Santos, 2017; Rosati et al., 2016). However, we did not perform these kinds of control comparisons in the current study because gaze following in this species is well-established (Bettle & Rosati, 2019; Drayton & Santos, 2017; Emery, 2000; Emery, Lorincz, Perrett, Oram, & Baker, 1997; Rosati et al., 2016; Rosati & Santos, 2017), and our main focus was on individual variation in gaze following. Moreover, both kinds of tasks might involve habituation over successive trials, but this means something different from a psychological perspective across the two tasks. Habituation is a core element of many looking time task designs, which are often aimed at either detecting dishabituation in response to a novel

This article is protected by copyright. All rights reserved.

Author Manuscript

change or more generally preferential allocation of attention (Margoni, Surian, & Baillargeon, 2022; Winters et al., 2015). In contrast, habituation over successive trials of gaze-following tasks is often taken as a marker of cognitive flexibility, as it indicates that animals control an adjust their responses rather than engaging in reflexive or automatic co-orienting repeatedly to the same gaze cue even when there is no actual target to detect (Rosati et al., 2016; Rosati & Hare, 2009; Schloegl, Kotrschal, & Bugnyar, 2007; Tomasello et al., 2001). Finally, it is important to note that while the socioemotional task used conspecific photographs, the gaze-following task used a human actor. While previous research shows that macaques robustly follow the gaze of both humans and conspecifics and appear to do so at similar rates (Ferrari et al., 2000; Rosati et al., 2016; Rosati & Santos, 2017; Teufel et al., 2010; Tomasello et al., 2001), there may be differences in how monkeys respond to some aspects of human versus conspecific social signals. Nonetheless, the fact that performance on these tasks were related in a given testing year suggests both are tapping into some shared social sensitivity.

Characterizing individual differences in cognition in the Cayo Santiago rhesus macaques allows for several new research questions moving forward. In particular, Cayo Santiago is a large population of monkeys living in naturalistic, freely-breeding social groups—but are also very well-habituated to human experimenters such that assessing cognitive performance using experimental techniques is possible as in the studies here. Moreover, demographics and social relationships in this population have been carefully documented over many generations (Ellis et al., 2019; Hernández-Pacheco et al., 2016; Maestripieri & Hoffman, 2012; Newman et al., 2023; Rawlins & Kessler, 1986). In contrast, animals in many other primate populations where cognitive tests can be performed

This article is protected by copyright. All rights reserved.

Author Manuscript

tend to live in smaller, human-controlled groups. This means that studies of the Cayo Santiago macaques are uniquely well-situated to connect cognitive traits to important biological outcomes such as health, longevity, and reproductive success. There is increasing evidence that social context impacts these kinds of outcomes in both humans and a variety of nonhuman animals (Dettmer & Chusyd, 2023; Lucas & Keller, 2020; Shively & Day, 2015; Snyder-Mackler et al., 2020). For example, individuals who are socially well-connected, have higher social status, or escape early life social adversity tend to have better health, longer lifespans, and greater fitness in several primate species (Morrison, Eckardt, Colchero, Vecellio, & Stoinski, 2021; Patterson et al., 2021; Silk, Alberts, & Altmann, 2003; Tung, Archie, Altmann, & Alberts, 2016). However, it is unclear whether or how underlying psychological propensities contribute to these kinds of effects.

The current work sets the stage for future work empirically testing these relationships between cognitive health, sociality and social status, and biological outcomes (Hernández-Pacheco et al., 2023). Given the findings here, one possibility is that some aspects of cognitive performance may be more strongly related to fitness and health metrics than others. For example, socioemotional responses—which show clear and consistent patterns of individual variation—may be more predictive of health, reproduction, and longevity than gaze following, which was not as consistent within individuals over time. That is, especially socioemotional responses may reflect a more stable trait that could be related to long-term biological outcomes and subject to the processes of natural selection. This may also be informative for understanding which cognitive abilities actually contribute to different metrics of social connectedness that are linked to aspects of fitness

This article is protected by copyright. All rights reserved.

Author Manuscript

such as longevity and reproductive success in primates (Silk, 2007). Disentangling how different cognitive traits are linked to such real-world outcomes is a critical next step for understanding the processes shaping the evolution of cognition across species.

Tables

Year of Test	*Socioemotional task*	*Gaze Task*	*Both tasks*
Year 1	170 (84F, 86 M)	169 (89F, 80M)	143 (74F, 69M)
Year 2	161 (79F, 82 M)	108 (55F, 53M)	93 (48F, 45M)
Both years	150 (75F, 75M)	102 (55F, 47M)	**85 (46F, 39M)**
At least one	181 (88F, 93M)	175 (89F, 86M)	
Full sample	**204 (101F, 103M)**		

Table 1: Participation by task and year. Breakdown shows total number of monkey and sex distribution (F = female, M = male) for the two tasks by year. We aimed to test all monkeys in both tasks at both timepoints but in some cases animals were not available (e.g., had died) or could not be located in appropriate testing locations (e.g., where we could place our apparatus). Note that socioemotional task participation here focuses on completion of the first two trials (with female conspecific photographs); some animals completed those trials but did not complete the final two trials (with male photographs) and so are included in only relevant analyses based on participation.

Experimental stimuli	*R*	*S.E.*	*N*	*C.I*	*p-value*
Female neutral expression*	0.178	0.080	150	0.026, 0.342	= 0.019
Female threat expression	0.277	0.079	150	0.117, 0.419	= 0.0004
Male neutral expression*	0.176	0.101	81	0.00, 0.393	= 0.077, n.s.
Male threat expression	0.356	0.098	81	0.171, 0.559	= 0.0008

Table 2: Repeatability in the socioemotional responses task. Long-term repeatability (R) estimates for looking time responses from cognitive experiments conducted approximately one year apart for the same individuals; each analysis also accounted for age and sex. *Indicates models with singular fit indicating little variance between individuals.

Trial	*R*	*S.E.*	*N*	*C.I*	*p-value*
Trial 1	0.157	0.077	102	0.00, 0.283	= 0.044
Trial 2	0.067	0.060	102	0.00, 0.213	= 0.209, n.s.
Trial 3	0.000	0.060	63	0.00, 0.205	= 1.0, n.s.
Trial 4*	0.243	0.1345	37	0.00, 0.952	= 0.089, n.s.

This article is protected by copyright. All rights reserved.

Author Manuscript

Table 3: Repeatability in the gaze following task. Long-term repeatability (R) estimates for co-orienting responses from experiments conducted approximately one year apart for the same individuals; each analysis also accounted for age and sex. All models had singular fit, indicating little variance between individuals. *Age and sex were not included in the model for trial 4, because separation of data prevented models from converging.

Acknowledgments

We thank Okiany Carillo Serrano, Natalia Camargo Peña, Phoebe Evans, Ismail Hazime, Yiyun Huang, and Nami Kaneko for assistance with data collection and coding. The authors also thank the Cayo Santiago Field Station staff including Alyssa Arre, Bianca Giura, Giselle Caraballo Cruz, and Nahiri Rivera Barreto for their support. This research was supported by grants from the National Institute on Aging (R21 20-0266), the National Science Foundation (2217812), and a Sloan Foundation Fellowship (FG-2019-12054). Alexis Diaz was supported by an NSF Graduate Research Fellowship (award 2141410). The Cayo Santiago Field Station is supported by a National Center for Research Resources CM-5-P40RR003640 award and an Office of Research Infrastructure Programs (ORIP) of the National Institutes of Health (NIH) through Grant Number P40OD012217 to the Caribbean Primate Research Center and the University of Puerto Rico.

Conflict of Interest Statement

The authors declare no conflict of interest.

References

Adolphs, R. (2008). Fear, faces, and the human amygdala. *Current Opinion in Neurobiology, 18*, 166-172.

Alberts, S. C., Altmann, J., Brockman, D. K., Cords, M., Fedigan, L. M., Pusey, A.,... Bronikowski, A. M. (2013). Reproductive aging patterns in primates reveal that humans are distinct. *Proceedings of the National Academy of Sciences, 110*, 13440-13445.

Almeling, L., Hammerschmidt, K., Senn-Reulen, H., Freund, A. M., & Fischer, J. (2016). Motivational shifts in aging monkeys amd the origins of social selectivity. *Current Biology, 26*, 1744-1749.

This article is protected by copyright. All rights reserved.

Author Manuscript

Almeling, L., Senn-Reulen, H., Hammerschmidt, K., Freund, A. M., & Fischer, J. (2017). Social interactions and activity patterns of old Barbary macaques: Further insights into the foundations of social selectivity. *American Journal of Primatology, 79*, e22711.

Alwall, N., Johansson, D., & Hansen, S. (2010). The gender difference in gaze-cueing: Associations with empathizing and systemizing. *Personality and Individual Differences, 49*, 729-732.

Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature, 554*, 364-367.

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology, 8*, 456.

Bates, D. (2010). The LME4 package: linear mixed-effects models using S4 classes. See http://www.R-project.org.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1-48.

Bayliss, A. P., di Pellegrino, G., & Tipper, S. P. (2005). Sex differences in eye gaze and symbolic cueing of attention. *The Quarterly Journal of Experimental Psychology, 58A*, 631-650.

Beran, M. J., & Hopkins, W. D. (2018). Self-control in chimpanzees relates to general intelligence. *Current Biology*.

Bettle, R., & Rosati, A. G. (2019). Flexible gaze following in rhesus monkeys. *Animal Cognition, 22*, 673-686.

Bettle, R., & Rosati, A. G. (2021). The evolutionary origins of natural pedagogy: Rhesus monkeys show sustained attention following nonsocial cues versus social communicative signals. *Developmental Science, 24*(1), e12987.

Bettle, R., & Rosati, A. G. (2022). Sensitivity to line-of-sight in tolerant versus despotic macaques (Macaca sylvanus and Macaca mulatta). *Journal of Comparative Psychology, 136*, 93-104.

Boesch, C., & Boesch, H. (1981). Sex differences in the use of natural hammers by wild chimpanzees: A preliminary report. *Journal of Human Evolution, 10*, 585-593.

Bogin, B., & Smith, B. H. (1996). Evolution of the human life cycle. *American Journal of Human Biology, 8*(6), 703-716.

Bohn, M., Eckert, J., Hanus, D., Lugauer, B., Holtmann, J., & Haun, D. B. M. (2023). Great ape cognition is structured by stable cognitive abilities and predicted by developmental conditions. *Nature Ecology and Evolution, 7*, 927–938.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2008). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution, 24*, 127-135.

Bond, A. B., Kamil, A. C., & Balda, R. P. (2007). Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (Gymnorhinus cyanocephalus, Nucifraga columbiana, Aphelocoma californica). *Journal of Comparative Psychlogy, 121*, 372-379.

Brent, L. J. N., Semple, S., Maclarnon, A., Ruiz-La, m., A., Gonzalez-Martinez, J., & Platt, M. L. (2014). Personality traits in rhesus macaques (Macaca mulatta) are

This article is protected by copyright. All rights reserved.

Author Manuscript

heritable but do not predict reproductive output. *International Journal of Primatology, 35*, 188-209.
Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*, 535-543.
Cantwell, A., Buckholtz, J. W., Atencia, R., & Rosati, A. G. (2022). The origins of cognitive flexibility in chimpanzees. *Developmental Science, 25*, e13266.
Carstensen, L. L., & Mikels, J. A. (2005). At the intersection of emotion and cognition: aging and the positivity effect. *Current Directions in Psychological Science, 14*, 117-121.
Carstensen, L. L., Turan, B., Scheibe, S., Ram, N., Ersner-Hershfield, H., Samanez-Larkin, G. R.,... Nesselroade, J. R. (2011). Emotional experience improves with age: Evidence based on over 10 years of experience sampling. *Psychology and Aging, 26*, 21-33.
Cauchoix, M. e. a. (2018). The repeatbility of cognitive perfomance: a meta-analysis. *Philosophical Transactions of the Royal Society B, 373*, 20170281.
Croston, R., Branch, C. L., Kozlovsky, D. Y., Dukas, R., & Pravosudov, V. V. (2015). Heritability and the evolution of cognitive traits. *Behavioral Ecology, 26*, 1447-1459.
D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3-to 6- month-olds. *Infant Behavior and Development, 20*, 569-572.
Dal Monte, O., Costa, V. D., Nobel, P. L., Murray, E. A., & Averbeck, B. B. (2015). Amygdala lesions in rhesus macaques decrease attention to threat. *Nature Communications, 6*, 10161.
De Petrillo, F., Bettle, R., & Rosati, A. G. (2022). Insights from matched species comparisons for understanding cognition in the wild. *Current Opinion in Behavioral Sciences, 45*, 101134.
De Petrillo, F., Nair, P., Cantwell, A., & Rosati, A. G. (2022). The evolution of cognitive control in lemurs. *Psychological Science, 33*, 1408-1422.
De Petrillo, F., & Rosati, A. G. (2019). Rhesus macaques use probabilities to predict future events. *Evolution and Human Behavior, 40*(5), 436-446.
De Petrillo, F., & Rosati, A. G. (2021). Variation in primate decision-making under uncertainty and the roots of human economic behaviour. *Philosophical Transactions of the Royal Society B 376*, 20190671.
Deaner, R. O., Shepherd, S. V., & Platt, M. L. (2007). Familiarity accentuates gaze cuing in women but not men. *Biology Letters, 2*, 64-67.
Dettmer, A. M., & Chusyd, D. E. (2023). Early life adversities and lifelong health outcomes: A review of the literature on large, social, long-lived nonhuman mammals. *Neuroscience & Biobehavioral Reviews, 152*, 105297.
Diaz, A. A., Steiner, U. K., Tuljapurka, S., Zuo, W., & Hernandez-Pacheco, R. (2023). Hurricanes affect diversification among individual life courses of a primate population. *Journal of Animal Ecology, 92*, 1404-1415.
Drayton, L. A., & Santos, L. R. (2015). A decade of theory of mind research on Cayo Santiago: Insights into rhesus macaque social cognition. *American Journal of Primatology, 78*, 106-116.
Drayton, L. A., & Santos, L. R. (2017). Do rhesus macaques, Macaca mulatta, understand what others know when gaze following? *Animal Behaviour, 134*, 193–199.

This article is protected by copyright. All rights reserved.

Author Manuscript

Dubuc, C., Allen, W. L., Cascio, J., Lee, D. S., Maestripieri, D., Petersdorf, M.,... Higham, J. P. (2016). Who cares? Experimental attention biases provide new insights into a mammalian sexual signal. *Behavioral Ecology, 27*, 68-74.

Ellis, S., Snyder-Mackler, N., Ruiz-Lambides, A., Platt, M. L., & Brent, L. J. N. (2019). Deconstructing sociality: The types of social connections that predict longevity in a group-living primate. *Proceedings B, 286*.

Emery, N. J. (2000). The eyes have it: the neuroethology, function, and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24*, 581-601.

Emery, N. J., Lorincz, E. N., Perrett, D. I., Oram, M. W., & Baker, C. I. (1997). Gaze following and joint attention in rhesus monkeys (Macaca mulatta). *Journal of Comparative Psychology, 111*, 286-293.

Ferrari, P. F., Coude, G., Gallese, V., & Fogassi, L. (2008). Having access to others' mind through gaze: The role of ontogenetic and learning processes in gaze-following behavior of macaques. *Social Neuroscience, 3*, 239-249.

Ferrari, P. F., Kohler, E., Fogassi, L., & Gallese, V. (2000). The ability to follow eye gaze and its emergence during development in macaque monkeys. *Proceedings of the National Academy of Sciences, 97*, 3997–14002.

Fichtel, C., Dinter, K., & Kappeler, P. M. (2020). The lemur baseline: how lemurs compare to monkeys and apes in the Primate Cognition Test Battery. *Peer J, 8*, e10025.

Fischer, J. (2017a). On the social life and motivation changes of aging monkeys. *Gerontology, 63*, 572-579.

Fischer, J. (2017b). On the social life and motivational changes of aging monkeys. *Gerontology, 63*, 572-579.

Ghazanfar, A., & Santos, L. R. (2004). Primate brains in the wild: the sensory bases for social interactions. *Nature Reviews Neuroscience, 5*, 603-616.

Gilby, I. C., Machanda, Z. P., O"Malley, R. C., Murray, C. M., Lonsdorf, E. V., Walker, K.,... Wrangham, R. W. (2017). Predation by female chimpanzees: Toward an understanding of sex differences in meat acquisition in the last common ancestor of Pan and Homo. *Journal of Human Evolution, 110*, 82-94.

Gilby, I. C., & Wrangham, R. W. (2008). Association patterns among wild chimpanzees (Pan troglodytes schweinfurthii) reflect sex differences in cooperation. *Behavioral Ecology and Sociobiology, 62*, 1831-1842.

Gomez, J. C. (2005). Species comparative studies and cognitive development. *Trends in Cognitive Sciences, 9*, 118-125.

Gothard, K. M., Battaglia, F. P., Erickson, C. A., Spitler, K. M., & Amaral, D. G. (2007). Neural responses to facial expression and face identity in the monkey amygdala. *Journal of Neurophysiology, 97*, 1671-1683.

Hartig, F. (2022). DHARMa: residual diagnostics for hierarchical (multi-level / mixed) regression models (Version R package version 0.4.6.). Retrieved from http://florianhartig.github.io/DHARMa/

Haux, L. M., Engelmann, J. M., Arslan, R. C., Hertwug, R., & Herrmann, E. (2022). Chimpanzees and human risk preferences show key similarlities. *Psychological Science, 34*.

Hernández-Pacheco, R., Delgado, D. L., Rawlins, R. G., Kessler, M. J., Ruiz-Lambides, A. V., Maldonado, E., & Sabat, A. M. (2016). Managing the Cayo Santiago

This article is protected by copyright. All rights reserved.

Author Manuscript

rhesus macaque population: The role of density. *American Journal of Primatology, 78*, 167–181.
Hernández-Pacheco, R., Steiner, U. K., Rosati, A. G., & Tuljapurka, S. (2023). Advancing methods for the biodemography of aging within social contexts. *Neuroscience and Biobehavioral Reviews, 153*, 105400.
Herrmann, E., Call, J., Hernadez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science, 317*, 1360-1366.
Herrmann, E., Hare, B., Call, J., & Tomasello, M. (2010). Differences in the cognitive skills of bonobos and chimpanzees. *PLoS One, 5*, e12438.
Herrmann, E., Hare, B., Cissewski, J., & Tomasello, M. (2011). A comparison of temperament in nonhuman apes and human infants. *Develomental Science, 14*, 1393-1405.
Herrmann, E., Hernández-Lloreda, M. V., Call, J., Hare, B., & Tomasello, M. (2010). The structure of individual differences in the cognitive abilities of children and chimpanzees. *Pschological Science, 21*, 102-110.
Higham, J. P., Hughes, K. D., Brent, L. J. N., Dubuc, C., Engelhardt, A., Heistermann, M.,... Stevens, M. (2011). Familiarity affects the assessment of female facial signals of fertility by free-ranging male rhesus macaques. *Proceedings of the Royal Society B, 278*, 3452–3458.
Hoffman, C. L., Higham, J. P., Mas-Rivera, A., Ayala, J. E., & Mastripieri, D. (2010). Terminal investment and senescence in rhesus macaques (Macaca mulatta) on Cayo Santiago *Behavioral Ecology, 21*, 972-978.
Hoffman, K. L., Gothard, K. M., Schmid, M. C., & Logothetis, N. K. (2007). Facial-expression and gaze-selective responses in the monkeys amygdala. *Current Biology, 17*, 766-772.
Huang, Y., Chang, H., Santos, L. R., & Rosati, A. G. (2024). Predictions about reward outcomes in rhesus monkeys. *Behavioral Neuroscience*.
Hughes, K. D., & Santos, L. R. (2012). Rotational displacement skills in the rhesus macaques (*Macaca mulatta*). *Journal of Comparative Psychology, 126*, 421-432.
Kominsky, J. F., Li, Y., & Carey, S. (2022). Infants' attributions of insides and animacy in causal interactions. *Cognitive Science, 46*, e13087.
Kuhn, G., Pagano, A., Maani, S., & Bunce, D. (2014). Age-related decline in the reflexive component of overt gaze following. *The Quarterly Journal of Experimental Psychology*.
Lacreuse, A., Schatz, K., Strazzullo, S., King, H. M., & Ready, R. (2013). Attentional biases and memory for emotional stimuli in men and male rhesus monkeys. *Animal Cognition, 16*, 861-871.
Leigh, S. R. (2004). Brain growth, life history, and cognition in primate and human evolution. *American Journal of Primatology, 62*, 139-164.
Leigh, S. R. (2012). Brain size growth and life history in human evolution. *Evolutionary Biology, 39*, 587-599.
Lenth, R. (2018). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.2.3. https://CRAN.R-project.org/package=emmeans.
Lo, S., & Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology, 6*, 1171.

This article is protected by copyright. All rights reserved.

Author Manuscript

Lonsdorf, E. V. (2005). Sex differences in the development of termite-fishing skills in the wild chimpanzees, Pan troglodytes schweinfurthii, of Gombe National Park, Tanzania. *Animal Behaviour, 70*, 676-683.

Lucas, E. R., & Keller, L. (2020). The co-evolution of longevity and social life. *Functional Ecology, 34*, 76-87.

Machanda, Z. P., Gilby, I. C., & Wrangham, R. W. (2013). Male–female association patterns among free-ranging chimpanzees (Pan troglodytes schweinfurthii). *International Journal of Primatology, 34*, 917-938.

Machanda, Z. P., & Rosati, A. G. (2020). Shifting sociality during primate ageing. *Philosophical Transactions of the Royal Society B, 375*, 20190620.

Maclean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C.,... Zhao, Y. (2014). The evolution of self-control. *Proceedings from the National Academy of Sciences*.

Maestripieri, D., & Hoffman, C. L. (2012). Behavior and social dynamics of rhesus macaques on Cayo Santiago. In Q. Wang (Ed.), *Bones, Genetics, and Behavior of Rhesus Macaques* (pp. 247–262). New York: Springer.

Margoni, F., Surian, L., & Baillargeon, R. (2022). The violation-of-expectation paradigm: A conceptual overview. *Psychological Review, 131*, 716-748.

Marticorena, D., Ruiz, A. M., Mukerji, C., Goddu, A., & Santos, L. R. (2011a). Monkeys represent others' knowledge but not their beliefs. *Developmental Science, 14*, 1406-1416.

Marticorena, D. C. W., Ruiz, A. M., Mukerji, C., Goddu, A., & Santos, L. R. (2011b). Monkeys represent others' knowledge but not their beliefs. *Developmental Science, 14*(6), 1406-1416.

Martin, A., & Santos, L. R. (2014). The origins of belief representation: Monkeys fail to automatically represent others' beliefs. *Cognition, 130*, 300-308.

Martin, A., & Santos, L. R. (2016). What cognitive representations support primate theory of mind? *Trends in Cognitive Sciences, 20*, 375-382.

Mason, D., Zajitschek, S., Anwer, H., O"Dea, R. E., Hesselson, D., & Nakagawa, S. (2021). Low repeatability of aversive learning in zebrafish (Danio rerio). *Journal of Experimental Biology, 224*, 240846.

Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences, 9*, 469-502.

Matsuzawa, T. (2007). Comparative cognitive development. *Developmental Science, 10*, 97-103.

Matsuzawa, T., Tomonaga, M., & Tanaka, M. (Eds.). (2006). *Cognitive development in chimpanzees*. Tokyo: Springer-Verlag.

McCrink, K., & Wynn, K. (2004). Large-number addition and dubtraction by 9-month-old infants. *Psychological Science, 15*(11), 776-781.

Morcillo, D. O., Steiner, U. K., Grayson, K. L., Ruiz-Lambides, A., & Hernandez-Pacheco, R. (2020). Hurricane-induced demographic changes in a non-human primate population. *Royal Society Open Science, 7*, 200173.

Morrison, R. E., Eckardt, W., Colchero, F., Vecellio, V., & Stoinski, T. S. (2021). Social groups buffer maternal loss in mountain gorillas. *elife, 10*, e62939.

This article is protected by copyright. All rights reserved.

Author Manuscript

Mundy, P., Block, J., Delgado, C., Pomares, Y., van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development, 78*, 938-954.

Newman, L. E., Testard, C., DeCasien, A. R., Chiou, K. L., Watowich, M. M., Janiak, M. C.,... Higham, J. P. (2023). The biology of aging in a social world: Insights from free-ranging rhesus macaques. *Neuroscience & Biobehavioral Reviews, 154*, 105424.

Oakes, L. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy, 22*, 436-469.

Okamoto, S., Tomonaga, M., Ishii, K., Kawai, N., Tanaka, M., & Matsuzawa, T. (2002). An infant chimpanzee (*Pan troglodytes*) follows human gaze. *Animal Cognition, 5*(2), 107-114.

Okamoto-Barth, S., Tomonaga, M., Tanaka, M., & Matsuzawa, T. (2008). Development of using experimenter-given cues in infant chimpanzees: longitudinal changes in behavior and cognitive development. *Developmental Science, 11*, 98-108.

Olineck, K. M., & Poulin-Dubois, D. (2009). Infants' understanding of intention from 10 to 14 months: Interrelations among violation of expectancy and imitation tasks. *Infant Behavior and Development, 32*, 404-415.

Patterson, S. K., Hinde, K., Bond, A. B., Trumble, B. C., Strum, S. C., & Silk, J. B. (2021). Effects of early life adversity on maternal effort and glucocorticoids in wild olive baboons. *Behavioral Ecology and Sociobiology, 75*, 114.

Pravosudov, V. V. (2022). Cognitive ecology in the wild — advances and challenges in avian cognition research. *Current Opinion in Behavioral Sciences, 45*, 101138.

Pritchard, D. J., Hurly, T. A., Tello-Ramos, M. C., & Healy, S. D. (2016). **Why study cognition in the wild (and how to test it)?** *Journal of the Experimental Analysis of Behavior, 105*, 41-55.

Pusey, A. E., Pintea, L., Wilson, M. L., Kamenya, S., & Goodall, J. (2007). The contribution of long-term research at Gombe National Park to chimpanzee conservation *Conservation Biology, 21*, 623-634.

R. Core Team. (2022). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rathke, E.-M., & Fischer, J. (2020). Differential ageing trajectories in motivation, inhibitory control and cognitive flexibility in Barbary macaques (Macaca sylvanus). *Philosophical Transactions of the Royal Society B, 375*, 20190617.

Rawlins, R. C., & Kessler, M. J. (Eds.). (1986). *The Cayo Santiago macaques: History, behavior and biology*. Albany, NY: State University of New York Press.

Rohrer, K. N., & Ferkin, M. H. (2020). Long-term repeatability and stability of three personality traits in meadow voles. *Ethology, 126*, 791-802.

Rosati, A. G., Arre, A. M., Platt, M. L., & Santos, L. R. (2016). Rhesus monkeys show human-like changes in gaze following across the lifespan. *Proceedings of the Royal Society B, 283*, 20160376.

Rosati, A. G., Arre, A. M., Platt, M. L., & Santos, L. R. (2018). Developmental shifts in social cognition: socioemotional biases across the lifespan in rhesus monkeys. *Behavioral Ecology and Sociobiology, 72*(163).

Rosati, A. G., DiNicola, L., & Buckholtz, J. W. (2018). Chimpanzee cooperation is fast and independent from self-control. *Psychological Science, 29*, 1832-1845.

This article is protected by copyright. All rights reserved.

Author Manuscript

Rosati, A. G., Emery Thompson, M., Atencia, R., & Buckholtz, J. W. (2023). Distinct developmental trajectories for risky and impulsive decision-making in chimpanzees. *Journal of Experimental Psychology: General*.

Rosati, A. G., Hagberg, L., Enigk, D. K., Otali, E., Emery Thompson, M., Muller, M. N.,... Machanda, Z. P. (2020). Social selectivity in aging wild chimpanzees. *Science, 370*, 473-476.

Rosati, A. G., & Hare, B. (2009). Looking past the model species: diversity in gaze-following skills across primates. *Current Opinion in Neurobiology, 19*, 45-51.

Rosati, A. G., Machanda, Z. P., & Slocombe, K. E. (2022). Cognition in the wild: understanding animal thought in its natural context. *Current Opinion in Behavioral Sciences, 47*, 101210.

Rosati, A. G., & Santos, L. R. (2017). Tolerant Barbary macaques maintain juvenile levels of social attention in old age, but despotic rhesus macaques do not. *Animal Behaviour, 130*, 199-207.

Rosati, A. G., Wobber, V., Hughes, K., & Santos, L. R. (2014). Comparative developmental psychology: How is human cognitive development unique? *Evolutionary Psychology, 12*, 448-473.

Sabbi, K. H., Emery Thompson, M., Machanda, Z. P., Otali, E., Wrangham, R. W., & Martin, M. N. (2021). Sex differences in early experience and the development of aggression in wild chimpanzees. *Proceedings of the National Academy of Sciences, 118*, e2017144118.

Santos, L. R., Nissen, A. G., & Ferrugia, J. A. (2006). Rhesus monkeys, Macaca mulatta, know what others can and cannot hear. *Animal Behavior, 71* 1175-1181.

Schloegl, C., Kotrschal, K., & Bugnyar, T. (2007). Gaze following in common ravens, Corvus corax: ontogeny and habituation. *Animal Behaviour, 74*, 769-778.

Schmitte, V., Pankau, B., & Fischer, J. (2012). Old world monkeys compare to apes in the primate cognition test battery. *PLoS One, 7*, e32024.

Schubiger, M. N., Fichtel, C., & Burkart, J. M. (2020). Validity of cognitive tests for non-human animals: Pitfalls and prospects. *Frontiers in Psychology, 11*, 1835.

Shaw, R. C., & Schmelz, M. (2017). Cognitive test batteries in animal cognition research: evaluating the past, present and future of comparative psychometrics. *Animal Cognition, 20*, 1003-1018.

Shepherd, S. V. (2010). Following gaze: gaze-following behaviors as a window into social cognition. *Frontiers in Integrative Neuroscience, 4*, 5.

Shively, C. A., & Day, S. M. (2015). Social inequalities in health in nonhuman primates. *Neurobiology of Stress, 1*, 156-163.

Silk, J. B. (2007). Social components of fitness in primate groups. *Science, 317*, 1347-1351.

Silk, J. B., Alberts, S. C., & Altmann, J. (2003). Social bonds of female baboons enhance infant survival. *Science, 302*, 1231-1234.

Slessor, G., Phillips, L. H., & Bull, R. (2008). Age-related declines in basic social perception: Evidence From tasks assessing eye-gaze processing. *Psychology and Aging, 23*, 812-822.

Slessor, G., Venturini, C., Bonny, E. J., Insch, P. M., Rokaszewicz, A., & Finnerty, A. N. (2014). Specificity of age-related differences in eye-gaze following: Evidence

This article is protected by copyright. All rights reserved.

Author Manuscript

from social and nonsocial stimuli. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*.

Snyder-Mackler, N., Burger, J. R., Gaydosh, L., Belsky, D. W., Noppert, G. A., Campos, F.,... Tung, J. (2020). Social determinants of health and survival in humans and other animals.Science. *Science, 368*, eaax9553.

Soha, J. A., Peters, S., Anderson, R. C., Searcy, W. A., & Nowicki, S. (2019). Performance on tests of cognitive ability is not repeatable across years in a songbird. *Animal Behaviour, 158*, 281–288.

Soravia, C., Ashton, B. J., Thornton, A., & Ridley, A. R. (2022). General cognitive performance declines with female age and is negatively related to fledging success in a wild bird. *Proceedings B, 289*, 20221748.

Spelke, E., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*, 605-632.

Testard, C., Larson, S. M., Watowich, M. M., Kaplinsky, C. H., Bernau, A., Faulder, M.,... Brent, L. J. N. (2021). Rhesus macaques build new social connections after a natural disaster. *Current Biology, 31*, 2299-2309.

Teufel, C., Gutmann, A., Pirow, R., & Fischer, J. (2010). Facial expressions modulate the ontogenetic trajectory of gaze-following among monkeys. *Developmental Science, 13*, 913-922.

Thompson Gonzalez, N., Machanda, Z., Otali, E., Muller, M. N., Enigk, D. K., Wrangam, R., & Emery Thompson, M. (2021). Age-related change in adult chimpanzee social network integration. *Evolution, Medicine, and Public Health, 9*, 448-459.

Thornton, A., & Lukas, D. (2012). Individual variation in cognitive performance: developmental and evolutionary perspectives. *Philosophical Transactions of the Royal Society B, 367*, 2773-2783.

Thornton, A., & Truskanov, N. (2022). The role of natural history in animal cognition. *Current Opinion in Behavioral Sciences, 46*, 101154.

Tigges, J., Gordon, T. P., McClure, H. M., Hall, E. C., & Peters, A. (1988). Survival rate and life span of rhesus monkeys at the Yerkes regional primate research center. *American Journal of Primatology, 15*(263-273).

Tkaczynski, P. J., Mielke, A., Samuni, L., Preis, A., Wittig, R. M., & Crockford, C. (2020). Long-term repeatability in social behaviours suggests stable social phenotypes in wild chimpanzees. *Royal Society Open Science, 7*, 200454.

Tomasello, M., & Carpenter, M. (2005). The emergence of social cognition in three young chimpanzees. *Monographs of the Society for Research in Child Development, 70*.

Tomasello, M., Hare, B., & Fogleman, T. (2001). The ontogeny of gaze following in chimpanzee, Pan troglodytes, and rhesus macaques, Macaca mulatta. *Animal Behaviour, 61*, 335-343.

Tomonaga, M., Tanaka, M., Matsuzawa, T., Myowa-Yamakoshi, M., & Kosugi, D. (2004). Development of social cognition in infant chimpanzees (Pan troglodytes): face recognition, smiling, gaze, and the lack of triadic interactions. *Japanese Psychological Research, 46*, 227-235.

Tung, J., Archie, E. A., Altmann, J., & Alberts, S. C. (2016). Cumulative early life adversity predicts longevity in wild baboons. *Nature Communications, 7*, 11181.

This article is protected by copyright. All rights reserved.

Author Manuscript

Völter, C. J., Reindl, E., Felsche, E., Civelek, Z., halen, A., Lugosi, Z.,... Seed, A. M. (2022). The structure of executive functions in preschool children and chimpanzees. *Scientific Reports, 12*, 6456.

Völter, C. J., Tinklenberg, B., Call, J., & Seed, A. M. (2018). Comparative psychometrics: establishing what differs is central to understanding what evolves. *Philosophical Transactions of the Royal Society B, 373*, 20170283.

Watowich, M. M., Chiou, K. L., Montague, M. J., Cayo Biobank Research Unit, Simons, N. D., Horvath, J. E.,... Snyder-Mackler, N. (2022). Natural disaster and immunological aging in a nonhuman primate. *Proceedings of the National Academy of Sciences, 119*, e2121663119.

Wellman, H. M. (2011). Developing a theory of mind. In U. Goswami (Ed.), *The Blackwell Handbook of Cognitive Development*: Wiley-Blackwell.

Wilson, A. J., Réale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A.,... Nussey, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology, 79*, 13-26.

Winters, S., Dubuc, C., & Higham, J. P. (2015). The looking time experimental paradigm in studies of animal visual perception and cognition. *Ethology, 121*, 625-640.

Wobber, V., Herrmann, E., Hare, B., Wrangham, R., & Tomasello, M. (2014). Differences in the early cognitive development of children and great apes. *Developmental Psychobiology, 56*, 547-573.

Wrangham, R. W. (2000). Why are male chimpanzees more gregarious than mothers? A scramble competition hypothesis. In P. M. K. (Ed.) (Ed.), *Primate Males: Causes and Consequences of Variation in Group Composition* (pp. 248-258). Cambridge: Cambridge University Press.

This article is protected by copyright. All rights reserved.