Rhesus monkeys show human-like changes in gaze following across the lifespan

Alexandra G. Rosati¹*, Alyssa M. Arre², Michael L. Platt³, Laurie R. Santos²

¹Department of Human Evolutionary Biology
Harvard University
Cambridge, MA 02138
*rosati@fas.harvard.edu

²Department of Psychology
Yale University
New Haven, CT 06511

³Department of Neuroscience
University of Pennsylvania
Philadelphia, PA 19104

Classification:
SOCIAL SCIENCES (Psychological and Cognitive Sciences)
BIOLOGICAL SCIENCES (Anthropology)

Key words:
Gaze following, social cognition, primates, cognitive development, life history, human evolution
Abstract

Gaze following is a foundational skill in human social-cognitive development. The emergence of this capacity scaffolds critical human-specific abilities such as theory of mind and communication, and disruptions in the typical development of gaze following in autism can have cascading effects on these skills. Nonhuman primates also follow others’ gaze, but less is known about how the cognitive mechanisms that support this behavior develop over the lifespan. Here we tested gaze following experimentally in 481 semi-free-ranging rhesus macaques (Macaca mulatta), the largest sample to date, ranging in age from infancy to old age. We found that monkeys’ propensity to follow gaze showed striking parallels with humans. First, macaque gaze following emerged in infancy and peaked in the juvenile period—indicating that younger monkeys were especially attuned to gaze direction information, as in humans. Furthermore, monkeys exhibited human-like sex-differences in gaze following after sexual maturity, as adult females were more likely to gaze follow than adult males. Finally, older monkeys showed reduced propensity of gaze following responses, as in human aging. Importantly, a second study indicated macaques exhibit similar baseline rates of looking upwards in a control condition, regardless of age. Our findings indicate that monkeys undergo robust ontogenetic shifts in gaze following across the lifespan that mirror those of humans, suggesting that our own species’ sensitivity to gaze cues builds upon ancestral cognitive mechanisms shared with other primates.
Significance Statement

Gaze following is a critical human social-cognitive skill that exhibits characteristic changes during early development and aging. Several theories propose that patterns of cognitive development in humans are unique amongst primates, but no previous study has examined whether nonhuman primates show these same changes in gaze following across the entire lifespan. We tested whether rhesus macaques, a highly social primate, show human-like changes in this capacity—despite exhibiting major differences in life history. We found that macaque ontogenetic trajectories parallel those seen in humans: gaze following emerges in infancy, is most flexible and robust in the juvenile period, and declines in old age. These findings indicate that human gaze following builds upon psychological mechanisms shared with other primates.
Introduction

Gaze following, or co-orienting with others, allows individuals to learn new information about their social or physical environment. Gaze following is a foundational skill in human social cognition, scaffolding the development of theory of mind and language (1-3). Human infants are especially attuned to other’s gaze direction, and use gaze as a cue signaling the communicative intention of others (4, 5). Disruptions in the typical early development of gaze following and social attention, as occurs in autism, highlight the critical importance of this skill in the acquisition of normal human social behavior and language (6-9). Yet human gaze following skills are not static after maturity, as there are continued and consistent developmental change in this skill throughout the life course. For example, human adults (but not infants) show gender differences in gaze following: women exhibit greater sensitivity to gaze cues than men (10-12). Furthermore, the propensity to follow gaze, as well as flexibility in social attention, declines in older adults (9, 13-15). Together, these observations indicate that human gaze following skills show characteristic changes across life history transitions including infancy and the juvenile period, the attainment of sexual maturity in adulthood, and finally reproductive senescence and aging.

Like humans, a variety of nonhuman primate species also follow gaze, but the mechanisms underlying this behavior may differ from humans (16, 17). For example, gaze following may emerge relatively late in some species’ development – supporting the hypothesis that nonhuman primates require a longer period of social interactions to bootstrap this skill – unlike the early emergence of gaze following in human infancy (18-24). Indeed, these differences in gaze following development are thought to play a critical role in differences in cognitive skills between humans and other species (25). Importantly, several theories link the
evolution of uniquely-human socio-cognitive abilities to the evolution of our species’ unique life history patterns (26, 27). In particular, human cognitive capacities are proposed to have co-evolved with human life history characteristics, such as an extended juvenile period (allowing the acquisition and refinement of skills) and increased longevity (allowing the exploitation of these skills). If human cognition and human life history evolved in tandem, then an important prediction is that other species with different life-history characteristics should exhibit different patterns of cognitive development. Consequently, studies of comparative cognitive development are a critical test of these theories. However, no previous study has examined nonhuman gaze following across the entire lifespan.

To address this gap, we studied the development of gaze following in a large population of semi-free-ranging rhesus macaques (Macaca mulatta) living on the island of Cayo Santiago (28). Rhesus monkeys are a highly-social Old World monkey, and this population lives under naturalistic conditions, in normal mixed-sex social groups with wide age variation from infancy to old age. Importantly, rhesus monkeys show robust similarities to humans in several social-cognitive capacities: they preferentially attend to socially-relevant information such as conspecific’s eyes and emotional expressions (29-32), and will even reason about the perceptual and goal states of human actors (33-36). Moreover, this species spontaneously follows the gaze of both conspecifics (37, 38) and humans (20), allowing us to use a human actor with precisely-controlled behavior to experimentally assay rhesus gaze following responses. Yet despite these cognitive commonalities, rhesus macaques also exhibit important differences in their life history characteristics compared to humans: macaques exhibit relatively faster growth and brain maturation during juvenility, no period of reproductive cessation, and a shorter absolute lifespan (39-41). Rhesus monkeys therefore show critical divergences in life history compared to humans.
in both early development and aging. Consequently, comparative patterns of gaze following in
rhesus macaques represent a strong test case for whether cognitive development and life history
are uniquely linked in humans, or whether our species’ patterns of cognitive change can be
shared with other primate species with varying life history characteristics.

Results

In Study 1, we tested the largest cross-sectional sample of nonhuman individuals to date
(n = 481) in an experimental gaze following task. Specifically, we examined patterns of gaze
following in four age cohorts based on life history transitions in this species: infants up to one
year old (completion of weaning); juveniles up to 5 years old (the onset of sexual maturity);
adults up to 15 years old; and older monkeys over 15 years old (monkeys in this population live
to approximately 20 years, only rarely exceeding 25 years). Monkeys completed one
experimental session consisting of up to four trials. In each trial, an experimenter captured the
monkey’s attention (for example, by calling their name or clapping; see Videos S1 and S2 for
examples from and older monkey and an infant, respectively). She then looked straight up for
10s (using both head and eye direction), and we measured whether the monkey looked in the
same direction during this time period.

Because such co-orienting behaviors can stem from different psychological processes—
ranging from more reflexive responses to other’s orientation to more complex inferences about
other’s perceptual experience and the nature of ‘seeing’ (16, 42, 43)—we examined two
additional features of monkeys’ behavior in order to disentangle the cognitive capacities
underlying their responses in our task. We first examined whether monkeys flexibly habituated
to the experimenter’s gaze over trials in the absence of a clear target, indicating that they
adjusted their response based on situational context (20) rather than just reflexively co-orienting
when another individual shifted their attention. Our second index was whether monkeys made
multiple independent looks to identify the (absent) target of gaze. In particular, we measured
whether monkeys looked up multiple times (e.g., looked up, looked away, and then looked up
again; see Video S3), similar to previous work examining whether primates “check back” with
the actor to assess their true line of sight (20, 44, 45).

We found that monkeys’ propensity to follow gaze showed major shifts across the
lifespan, much like in humans. Whereas 36.2 ± 7.1% of infants looked up on their first trial, 64.4
± 3.9% of juveniles looked up, 48.3 ± 3.2% of adults, and only 24.4 ± 6.5% of older monkeys
(see Figure 1). That is, gaze following showed a nonlinear pattern by age: it emerged in infancy,
peaked in juveniles, and then declined across adulthood and old age. We used generalized linear
mixed models (GLMM) in R to examine gaze following as a binomial outcome, while also
accounting for within-subject repeated measures across trials with unequal repeats (46) as our
free-ranging subjects completed different total numbers of trials. We compared models using
likelihood ratio tests (see SI for analysis details). We first fitted a base model that included
random subject intercepts to account for repeated measures, and trial number as a covariate to
account for any within-session shifts in gaze following. To assess the importance of life history
stage on gaze following, we then added age cohort in the second model. Including cohort as a
predictor significantly improved fit as compared to the base model \( \chi^2 = 30.10, \text{df} = 3, p <
0.001 \); pairwise comparisons revealed that juveniles followed gaze more than all other groups,
and younger adults further followed gaze more than older adults [Tukey tests: p < 0.01 for all
significant cases]. In the third model, we finally added sex as a factor to assess if the monkeys
showed human-like sex differences in responses. This further increased model fit compared to
the second model [$\chi^2 = 5.64, \text{df} = 1, p < 0.05$], revealing that males were overall less likely to follow gaze than females were (see Table 1 for parameters from the full model).

We next examined the factors predicting gaze following responses within each cohort. We were particularly interested in whether different cohorts *habituated* by showing reductions in their propensity to look up across trials, our first index of the cognitive processes supporting monkey gaze following (see Figure 1). We also wished to examine when the sex difference in monkey gaze following emerged. To do so, we used GLMM to model the importance of *trial number*, *sex*, and *age* within each cohort separately (see SI for details and tables S1-S4 for cohort models). We found that different factors predicted gaze following responses across our age cohorts. Infants (n = 47) did not show habituation or sex differences in responses, but model fit was improved by including age [$\chi^2 = 17.89, \text{df} = 1, p < 0.001$]. Indeed, the youngest monkey in our sample that followed gaze was 5.5 months old, indicating that gaze following first emerges during the second half of the first year of life and then increases towards the juvenile period. In contrast, in juveniles (n = 149) the inclusion of only trial number improved model fit [$\chi^2 = 15.14, \text{df} = 1, p < 0.001$]. That is, juveniles were less likely to look up over subsequent trials, indicating that flexible habituation to repeated gazes emerges in the juvenile period. For adult macaques (n = 240), model fit was improved by including both trial number [$\chi^2 = 13.76, \text{df} = 1, p < 0.001$], and also sex as a predictor [$\chi^2 = 6.92, \text{df} = 1, p < 0.01$]. Thus, adults also showed flexible habituation, and human-like sex differences in social attention first emerge in macaques following sexual maturity. Finally, none of these predictors improved model fit in older adults (n = 45), indicating that flexible control over gaze following responses declined during monkey aging.
Next we examined the number of *discrete looks* that monkeys made, our second index of the cognitive processes supporting gaze following (see Figure 2). We predicted that if monkeys exhibit enhanced sensitivity to gaze cues in infancy, then infants should make the most number of discrete looks in an effort to locate the (absent) target. We therefore analyzed trial-one behavior of monkeys who followed gaze on that trial (e.g., before they could receive any feedback that there was in fact no target). Within the subset of monkeys who followed gaze (n = 240), we found that the number of discrete looks upwards was negatively correlated with age \[ r_p = -0.17, \ p < 0.01 \]. Whereas infants made an average of 2.0 ± 0.31 looks, this declined to 1.36 ± 0.20 in older adults (see Figure 2). One possibility is that infants were making the most robust attempts to locate the experimenter’s target when they did not initially identify it, whereas older monkeys have learned to inhibit these repeated looking responses.

We then examined several alternative explanations for this pattern of results. First, we tested whether differences in habituation responses across cohorts might be due to age differences in the total number of trials that monkeys completed. We found that infants completed slightly fewer trials than adults \[ F_{3,477} = 2.79, \ p < 0.05; \ Tukey \ tests \ indicated \ only \ infants \ and \ adults \ differed \ in \ pairwise \ comparisons \], but that differences in trial completion cannot account for our overall pattern of results (see SI for details). Second, we considered whether cohorts differed in their baseline tendency to look up, as juveniles may have looked up more often than other cohorts in our experiment simply because they made more baseline upward looks in general. To examine this, we assessed whether there were differences in the latency or duration of the juveniles’ looks, because baseline looking that did not occur in response to the experimenter’s actions should (1) be randomly distributed across the trial, rather than occurring soon after the experimenter looked up; and (2) last a shorter duration, as monkeys
would not be seeking to locate any target in that location. We therefore examined trial one
performance for individuals who did look up (see SI for details), but we found no significant
differences in either the cohort’s latency to look up \( [F_{3,236} = 1.44, p > 0.23, \text{n.s.}] \) or duration of
gaze \( [F_{3,236} = 1.95, p > 0.12, \text{n.s.}] \). That is, individuals who looked up tended to show similar
latencies and durations of looking, regardless of their cohort. Thus, these results do not support
the claim that differences in baseline looking rates across cohorts account for our main findings.

In our second study, we then specifically tested whether our results could be due to
differences in baseline rates of looking up. Although there is robust evidence from previous work
(20, 38) that rhesus monkeys follow the specific gaze direction of an actor (e.g., look up more in
the test situation than in a control condition), our first study did not directly measure baseline
looking rates. Study 2 therefore retested a subset of 80 juveniles, adults, and older adults using
an identical method to Study 1, except that the experimenter alternated the direction of her gaze
across trials (straight up versus straight down). This method allowed us to assess whether the
monkeys were more likely to look up on \textit{Up trials} after the experimenter looked in that specific
direction, compared to \textit{Down trials} where her behavior was identical but she looked in different
direction. This therefore provided a measure of baseline rates of upward gazing in this context.

We found that overall monkeys looked up on \( 40.4 \pm 4.1 \% \) of \textit{Up trials}, but only \( 14.1 \pm 2.9 \% \) of \textit{Down trials} (see Figure 3). We again used GLMMs to examine what factors predicted
whether individuals would look up. We first fitted a base model that included random \textit{subject}
intercepts, \textit{condition order} (subject’s random assignment to experience either an up or down trial
first), and \textit{trial number} as a covariate to account for the within-session habituation. In the second
model, we then added \textit{age} as a covariate to examine if propensity to gaze-follow declined (as
infants were not included, age was predicted to have a linear effect on gaze following in this
sample). We found that gaze following indeed declined with age, as this predictor improved model fit compared to the base model \( \chi^2 = 16.72, \text{df} = 1, p < 0.001 \). Finally, we tested the efficacy of our experimental manipulation by adding trial condition (up or down) as a predictor. This further improved model fit \( \chi^2 = 34.98, \text{df} = 1, p < 0.001 \), indicating that the monkeys were more likely to look up after the experimenter looked up, compared to the control condition where she looked in a different direction (see Table 2 for parameters from the full model).

To specifically assess whether the age effects seen in Study 1 could be accounted for by difference in baseline rates of looking upwards across ages, we next examined how responses on Up and Down trials separately related to age (see SI for details). We found that the subject’s age was a significant predictor of responses on Up trials, as including age increased model fit \( \chi^2 = 14.74, \text{df} = 1, p < 0.001 \). In particular, monkeys’ gaze following responses declined with age, replicating the basic results from Study 1. In contrast, including age did not increase model fit for Down trials \( \chi^2 = 1.58, \text{df} = 1, p > 0.20, \text{n.s.} \). Study 2 therefore replicated the main results from Study 1—showing that rates of gaze-following declined between the juvenile period and old age—but also confirmed that baseline rates of looking upwards remained at similar low rates across all ages. Consequently, our main developmental results from Study 1 reflect differences in gaze following responses, not baseline reactivity.

**Discussion**

Several proposals argue that the evolution of human life history and the evolution of human cognition are intertwined: our relatively long juvenile period allows the enhancement and elaboration of sophisticated cognitive skills, and our long lifespan allows for the exploitation of the cognitive skills we have acquired (26, 27). Studies of comparative cognitive development in
species that vary in life-history patterns are therefore critical empirical tests of this hypothesis. Our findings endorse the hypothesis that gaze following in our own species builds upon ancestral mechanisms shared with other primates. Thus, ontogenetic trajectories for at least some critical socio-cognitive capacities may be relatively conserved—even across species with fairly divergent life history characteristics compared to humans—as rhesus macaques undergo human-like developmental changes in social attention. Study 1 found that rhesus macaque gaze following emerged in infancy—in contrast to evidence from other primate species showing a relatively later emergence (18, 21). Human infants and toddlers use gaze cues as a window into other people’s mental states and communicative intentions (1, 3, 4), and we found that younger monkeys were also especially attuned to others’ gaze. For example, younger monkeys made the most discrete looks in the direction of the (absent) target, showing increasing responsiveness to gaze direction. Moreover, juvenile monkeys showed the highest propensity to follow gaze—and they flexibly controlled their responses by habituating to repeated gazing across trials.

Our approach of examining macaques across the lifespan further allowed us for the first time to test patterns of gaze following in later life, specifically sex differences in responses and changes during aging. In humans, adult women are more responsive to gaze cues than are men (10, 11). We found that adult rhesus monkeys exhibited this same sex difference in responses in Study 1, a finding that has not been consistently detected in nonhuman primates (16, 47, 48). Gender differences in human cognition may stem from diverse causes—ranging from socialization and cultural norms of behavior, to biological variation in underlying neurobiological systems—so such comparative data from species that lack human-like socialization practices can help disentangle these different influences on cognition. We further found that older monkeys show declines in gaze following, which is the first empirical
demonstration that monkeys and humans show similar patterns of change in social attention during healthy aging. Older adult humans show reductions in gaze sensitivity and flexibility in gaze following responses compared to younger adults (14, 15), and older monkeys likewise were less likely to follow gaze and did not exhibit the same patterns of flexible habituations seen in juveniles and younger adults. The results from Study 2 indicate that these age-related changes were due specifically to changes in gaze following propensity, not baseline differences in tendency to look upwards more generally. Together, these results suggest that the ontogenetic changes seen in human social attention across the life course are surprisingly similar to those seen in rhesus macaques as well.

Given that rhesus macaques show similar patterns of social attention development despite exhibiting quite different life-history characteristics than humans, what accounts for the important differences in human and macaque social cognition? One possibility is that humans and other primates do not differ in their pace of gaze following development, but rather in the developmental patterns that link the emergence of different skills across ontogeny (25). In humans, several skills emerge after and build upon early gaze following skills—including joint attention, complex theory of mind, and communication and language (1-3). Studies of developmental disorders where normal patterns of gaze following and social attention are disrupted, such as autism, further demonstrate the importance of these early-emerging skills for the development of normal human social behavior (6-9). While rhesus macaques (and other nonhuman primates) also follow gaze, they do not communicate in human-like ways or acquire other mentalizing skills that emerge in human development, such as understanding false beliefs (36, 49). Indeed, evidence from ape ontogeny suggests that even socio-cognitive capacities that do emerge in primates—such as understanding others’ goals or social learning—are not related
to earlier-emerging gaze following capacities, as they are in human children (23). This suggests that the human developmental link between early-emerging gaze following and other social-cognitive capacities is not necessarily shared with other species. Thus, while humans and other primates may show a similar rate of cognitive change in gaze following, this capacity may spur the development of other social cognitive skills in humans alone.

Materials and Methods

Subjects

Study 1 included 481 rhesus macaques (219 males and 262 females) ranging from two weeks to 28.5 years old: 47 infants, 149 juveniles, 240 adults, and 45 older monkeys. The Cayo Santiago population in Puerto Rico consists of over 1200 free-ranging, individually-identifiable monkeys. Monkeys had to successfully complete at least one trial to be included; additional monkeys were tested but excluded because they ran away without being identified, were judged to be not looking at the start of their only trial during coding, or were not in view for the full 10s of the trial (see SI for details). In Study 2, we re-tested 80 individuals (39 males and 41 females) ranging from 1.5 to 22 years old. Monkeys had to successfully complete at least the first two trials (one Up and one Down trial) due to the within-subjects condition comparison; additional monkeys were excluded because they did not successfully complete at least two trials (see SI).

Procedure

Two experimenters approached a calmly sitting monkey (standing 1-2 m away). Experiment 1 (E1), the actor, first attracted the monkey’s attention to her face (by calling “monkey” and/or clapping her hands). In Study 1 she then said “now” and looked directly up
Experimenter 2 (E2), the cameraperson, stood next to E1 and filmed the monkey’s face; she said “stop” after the 10s trial had concluded. E1 did not look at an actual target, but to make the experimenter’s action appear plausible we tested monkeys sitting in the vicinity of a tree where it was plausible that she was looking up at something. We consequently refrained from testing in locations when another monkey was present in the tree above the subject, to avoid any possible visual and auditory confounds. Monkeys completed up to four identical trials if they did not voluntarily leave the vicinity (as monkeys were free-ranging on the island during the test). If the monkey ran away or moved to an inaccessible location, E1 (who was blind to the monkey’s previous responses, as she had been looking up) decided whether to end the session. Study 2 used the same general procedure, but alternated whether E1 looked straight up or straight down across the four trials. We counterbalanced whether monkeys tested in Study 2 first experienced an *Up trial* or a *Down trial*. Monkeys were never tested on the same day as their participation in Study 1 (typically at least two weeks later).

**Coding**

Two independent coders scored the monkeys’ responses on all trials: a primary coder (who scored the entire dataset) and a second reliability coder (two reliability coders each scored a subset of the data, which was collected over a two year period). We first clipped out individual trials from longer session videos, and then randomized the order of trials (assigning a new, random trial ID) to blind coders to study number, trial number, and condition. Each coder independently identified the start of the trial (e.g., when E1 said “now”) and examined the subsequent 10s period frame-by-frame to judge the monkey’s response. For each trial, we coded (1) whether the monkey ever looked upwards (using either their entire head or eyes only) during
the 10s period; (2) the total duration of looking upwards; (3) the monkey’s latency to look up; and (4) the total number of discrete looks (e.g., looking up, looking away, and then looking up again). The primary and reliability coders had high reliability for all measures; for example, on their subset compared to the primary coder’s scores, both reliability coders agreed with the primary coder at Kappa > 0.92 for whether the monkey looked up on a given trial, and showed correlations of $r_p > 0.95$ for the total duration of looking across trials (see SI for all coding and reliability details).
Acknowledgments

We thank Lindsey Jones and Samantha Monier for assistance with data collection, and
Linda Chang and Janelle Gagnon for assistance with data coding. The authors are also grateful to
Angelina Ruiz Lambides, Nahiri Rivera Barreto, and Giselle Caraballo Cruz for their research
support at the Cayo Santiago Field Station. Research methods were approved by the Cayo
Santiago IACUC and the Yale IACUC, and adhered to guidelines for animal research at this site.
This research was supported by the NIMH (R01MH096875) and an NCRR grant CM-5-
P40RR003640-13 award to the Caribbean Primate Research Center the University of Puerto Rico-Medical Sciences Campus. LRS was supported by a James S. McDonnell Foundation Award (#220020242), and Yale University.
### Table 1: Factors influencing propensity to follow gaze in macaques’ development (Study 1). Predictors from the full (best fit) model. *Trial number* (1-4) was included as a covariate across models, and we added *cohort* and *sex* in successive models to test their importance.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial number</td>
<td>-0.308</td>
<td>0.062</td>
<td>-4.979</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Sex <em>(Female baseline)</em></td>
<td>-0.402</td>
<td>0.170</td>
<td>-2.370</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>Cohort: Juveniles v. infants</td>
<td>1.017</td>
<td>0.328</td>
<td>3.097</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Cohort: Adults v. infants</td>
<td>0.436</td>
<td>0.311</td>
<td>1.404</td>
<td>0.18</td>
</tr>
<tr>
<td>Cohort: Older adults v. infants</td>
<td>-0.701</td>
<td>0.411</td>
<td>-1.706</td>
<td>0.18</td>
</tr>
<tr>
<td>Cohort: Adults v. juveniles</td>
<td>-0.581</td>
<td>0.192</td>
<td>-3.034</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Cohort: Older adults v. juveniles</td>
<td>-1.718</td>
<td>0.338</td>
<td>-5.086</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Cohort: Older adults v. adults</td>
<td>-1.137</td>
<td>0.320</td>
<td>-3.557</td>
<td>&lt; 0.005</td>
</tr>
</tbody>
</table>

### Table 2: Factors influencing propensity to follow gaze in direction control study (Study 2). Predictors from the full (best fit) model. *Trial number* (1-4) as a covariate and the factor *condition order* (direction of experimenter’s look on the first trial) were included as predictors across models, and we added *age* and *trial condition* in successive models to test their importance.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial number</td>
<td>0.038</td>
<td>0.156</td>
<td>0.247</td>
<td>0.80</td>
</tr>
<tr>
<td>Cond. Order <em>(Down First baseline)</em></td>
<td>-0.900</td>
<td>0.459</td>
<td>-1.959</td>
<td>&lt; 0.051</td>
</tr>
<tr>
<td>Age</td>
<td>-0.224</td>
<td>0.061</td>
<td>-3.665</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Trial Cond. <em>(Down Trial baseline)</em></td>
<td>2.037</td>
<td>0.401</td>
<td>5.085</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
Figures

Figure 1: Gaze following responses across the lifespan (Study 1). Monkeys completed an experimental gaze following task in which a human attracted their attention and then looked up (for up to 4 trials). Juveniles and adults exhibited flexible habituation in response to repeated trials. Error bars indicate SE.

Figure 2: Number of discrete looks during gaze following (Study 1). Monkeys could make multiple discrete looks in order to locate the (absent) target of the experimenter’s gaze. When they looked up, younger monkeys made more discrete looks than did older monkeys. Error bars indicate SE.
Figure 3: Responses to different gaze directions (Study 2). The control study assessed how monkeys responded to the experimenter looking up versus down (as a measure of baseline upward looking). While gaze following responses declined with age in the Up trials, monkeys of all ages exhibit similar low rates of baseline upward looks. Error bars indicate SE.
References


Supplemental Information
Rhesus monkeys show human-like changed in gaze following across the lifespan
Alexandra G. Rosati, Alyssa M. Arre, Michael L. Platt, Laurie R. Santos

Subjects
We tested semi-free-ranging monkeys from the Cayo Santiago population. These monkeys live in natural social groups, are provisioned with monkey chow at feeding corrals each day (in addition to access to plants growing on the island), and have ad-lib access to water. Monkeys are identifiable by unique combinations of tattoos and ear notches, and are familiar with human experimenters. Infants younger than a year do not yet have a tattoo, but their birthdates and sex are recorded in the census in association with their mother. Consequently, to test infants without tattoos, we identified babies in close proximity to an identifiable female (e.g., clinging to a female for an extended period or nursing from her) and then identified that female’s baby in the census.

General Methods
As reported in the main text, two experimenters approached a calmly sitting monkey (standing 1-2 m away from the monkey). Experimenter 1 (E1) attracted the monkey’s attention to her face; once the monkey was looking she said “now” and looked directly up (Figure S1). Experimenter 2 (E2) filmed the monkey’s face for the 10s after E1 looked up.

![Figure S1: Experimental demonstration.](image)

(a) E1 stood 1-2m away from a calmly-sitting monkey and attracted their attention. (b) Once the monkey looked at her, she looked straight up.

Subject Exclusions
In Study 1, additional monkeys were approached by the experimenters for testing but were not included in the analyses because they ran away before being successfully identified (n = 12). Monkeys had to complete at least one complete trial where they were appropriately watching the experimenter when she called “now” and looked up, and they remained visible throughout the 10s trial (e.g., a monkey who ran out of sight before the 10s was complete could
not be properly coded for gaze-following). An additional 5 monkeys were excluded because they were originally tested on only one trial, and subsequently the blind coders scored that to be a bad trial (e.g., monkey not looking when E1 said “now” or the monkey ran out of shot). If a given trial was scored as bad but that individual successfully completed other trials, we analyzed only those successful trials (an additional 27 trials were excluded from monkeys that were otherwise included in the dataset; as a total 1397 trials were included in the analyses, this means we excluded a minority of only 1.8% of trials). If monkeys were tested in more than one session (for example, because they were identified only after the test and it turned out they had been tested previously), we used only the first session where they successfully completed at least one trial for analyses.

In Study 2, we re-tested monkeys that had previously completed Study 1. Monkeys had to successfully complete at least the first two trials (e.g., one Up and one Down trial) to be included in the study due to the within-subjects design. Consequently, additional monkeys were approached by the experimenters for testing but were not included in the analyses because they only completed one trial (n = 10), or because the coders scored that the monkey was not looking when the primary experimenter called “now” on one of the first two trials (n = 4). An additional 2 trials (from the second half of the test) were excluded from monkeys who successfully completed the first two trials and were therefore included in the final sample.

Coding
We first clipped out individual trials (from longer session videos) using the program MPEG Streamclip. Each trial was clipped starting a few second before E2 called “now” and ending after E2 said “stop.” Each resultant trial clip was assigned a new random identifying number that randomized the order of trials across all monkeys tested in Study 1 and Study 2. This allowed us to code each trial blind to trial number, as well as the total number of trials each monkey completed. The second phase of data collection (in 2014-15) randomized trials from Study 1 and Study 2, and therefore coding was also blind to study and condition (as Study 2 included both trials where E1 looked up, and control trials where she looked down).

All trials were scored by two independent coders. A primary coder examined all trials from the entire study. A first reliability coder scored the first half of sessions (data collected in 2013), and a second reliability coder scored the data collected in the second half (2014-2015). For each trial clip, coders independently identified the start of the word “now” and examined each trial frame-by-frame (30 frames = 1s) for the 10s after this start time. We coded: (1) whether the monkey ever looked straight upwards (using either their entire head or eyes only) in the ten second period of the trial; (2) the total duration of time the monkey spent looking up; (3) latency to initially look up; and (4) total number of discrete looks the monkey made (e.g., looking up, looking back down or at the experimenter, and then gazing up again; see Video S2).

Reliability between the primary coder and the reliability coders were high for all of these measures. For the first reliability coder, agreement for whether or not the monkey looked up was Kappa = 0.92; correlation of total time spent looking up was r_p = 0.96; correlation for the total number of discrete looks was r_p = 0.87; and correlation for latency to look up was r_p = 0.92. For the second reliability coder, agreement for whether or not the monkey looked up was Kappa = 0.94; correlation of total time spent looking up was r_p = 0.95; correlation for the total number of discrete looks was r_p = 0.92; and correlation for latency to look up was r_p = 0.90.
Statistical Analysis

For our main analyses, we used the `glmer` function from the `LME4` software package (Bates 2010) in R (R Development Core Team 2014) to model whether monkeys followed gaze. We fit binomial models a logit link function using maximum likelihood, including random subject intercepts to account for repeated trials within subjects (Baayen 2008). We conducted post-hoc Tukey comparisons of model factors using the `glht` function in the package `multcomp` (Hothorn et al 2008), and compared the fit of different models using likelihood ratio tests (Bolker et al 2008).

**Study 1 Supplementary Results**

**Analysis of habituation within cohorts**

As reported in the main text, we examined what factors predicted gaze following responses within each cohort separately. For each separate cohort, we used GLMM modeling to test the importance of trial number, sex, and age (as a linear predictor within each cohort) by sequentially adding predictors to a base model including only subject as a random factor, and then retained predictors that improved model fit. This allowed us to determine the specific factors that predicted gaze following within each cohort separately.

(1) **Gaze following in infants (n = 47).** In infants, initial models adding trial and sex did not improve model fit, compared to a base model with only subjects as a random factor [trial model versus base model: $\chi^2 = 1.29$, df = 1, p = 0.26, n.s.; sex model versus base model: $\chi^2 = 0.65$, df = 1, p = 0.42, n.s.]. This indicates the infants did not show habituation across trials, and also did not differ by sex. However, adding in age as a linear predictor did improve fit [age model versus base model: $\chi^2 = 17.89$, df = 1, p < 0.001], indicating that propensity to follow gaze increased as a function of age in this cohort. The parameters from a full model including all predictors are reported in Table S1.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial number (covariate)</td>
<td>-0.470</td>
<td>0.304</td>
<td>-1.546</td>
<td>0.12</td>
</tr>
<tr>
<td>Sex (female baseline)</td>
<td>-0.841</td>
<td>1.106</td>
<td>-0.760</td>
<td>0.45</td>
</tr>
<tr>
<td>Age (covariate)</td>
<td>7.455</td>
<td>2.576</td>
<td>2.894</td>
<td>&lt; 0.005</td>
</tr>
</tbody>
</table>

**Table S1:** Factors influencing propensity to follow gaze in macaque infants (Study 1). Parameters from the full model; the best fit model included only age.

(2) **Gaze following in juveniles (n = 149).** In juveniles, including trial increased model fit [trial model versus base model: $\chi^2 = 15.14$, df = 1, p < 0.001], indicating that this age cohort showed habituation across trials. In subsequent models we then added sex and age, but neither of these terms further improved model fit [sex-model versus trial-only model: $\chi^2 = 0.01$, df = 1, p = 0.95, n.s.; age-model versus trial-only model: $\chi^2 = 1.13$, df = 1, p = 0.29, n.s.]. Thus, within juveniles there was not a sex difference in responses, and no major shifts by age within this cohort. The parameters from a full model including all predictors are reported in Table S2.
Table S2: Factors influencing propensity to follow gaze in macaque juveniles (Study 1). Parameters from the full model; the best fit model included only trial number.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial number (covariate)</td>
<td>-0.421</td>
<td>0.112</td>
<td>-3.769</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Sex (female baseline)</td>
<td>0.058</td>
<td>0.290</td>
<td>0.200</td>
<td>0.84</td>
</tr>
<tr>
<td>Age (covariate)</td>
<td>-0.134</td>
<td>0.124</td>
<td>-1.078</td>
<td>0.28</td>
</tr>
</tbody>
</table>

(3) Gaze following in adults (n = 240). In adults, including trial increased model fit as in juveniles [trial model versus base model: \( \chi^2 = 13.76, \text{df} = 1, p < 0.001 \)], indicating that this age cohort also showed habituation across trials. In addition, adding sex also improved model fit [sex-model versus trial-only model: \( \chi^2 = 6.92, \text{df} = 1, p < 0.01 \)]. However, adding age did not improve model fit [age-model versus sex and trial model: \( \chi^2 = 0.40, \text{df} = 1, p = 0.52, \text{n.s.} \)]. Thus, adults showed flexible habituation like juveniles, but sex differences in responses also emerged in this age group. The parameters from a full model are reported in Table S3.

Table S3: Factors influencing propensity to follow gaze in macaque adults (Study 1). Parameters from the full model; the best fit model included trial number and sex.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial number (covariate)</td>
<td>-0.287</td>
<td>0.082</td>
<td>-3.495</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Sex (female baseline)</td>
<td>-0.604</td>
<td>0.226</td>
<td>-2.670</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Age (covariate)</td>
<td>-0.028</td>
<td>0.045</td>
<td>-0.632</td>
<td>0.53</td>
</tr>
</tbody>
</table>

(4) Gaze following in older adults (n = 45). In older adults, in contrast, none of these predictors improved model fit compared to the base model [trial model versus base model: \( \chi^2 = 0.30, \text{df} = 1, p = 0.58, \text{n.s.} \); sex-model versus base-model: \( \chi^2 = 1.91, \text{df} = 1, p = 0.17, \text{n.s.} \); age-model versus base-model: \( \chi^2 = 0.15, \text{df} = 1, p = 0.70, \text{n.s.} \)]. Thus, older adults showed less flexible control over the responses as evidenced by the lack of habituation across trials. The parameters from a full model including all predictors are reported in Table S4.

Table S4: Factors influencing propensity to follow gaze in macaque older adults (Study 1). Parameters from the full model; no predictors improved model fit compared to the base model.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial number (covariate)</td>
<td>-0.108</td>
<td>0.202</td>
<td>-0.536</td>
<td>0.59</td>
</tr>
<tr>
<td>Sex (female baseline)</td>
<td>-0.833</td>
<td>0.646</td>
<td>-1.291</td>
<td>0.20</td>
</tr>
<tr>
<td>Age (covariate)</td>
<td>0.004</td>
<td>0.092</td>
<td>0.041</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Trial Completion

One possible explanation for our main results is that the differences in habituation responses across cohorts were due to age differences in the total number of trials that monkeys completed. We compared number of completed trials using an ANOVA with age cohort as a between-subjects factor. In fact, infants completed fewer total trials than adults [\( F_{3,477} = 2.79, p < 0.05 \); posthoc Tukey test revealed only infant-adult comparison \( p < 0.05 \); infant mean 2.51 ± 0.18 trials; juveniles 2.81 ± 0.10; adults 3.02 ± 0.08; older adults 3.00 ± 0.19 trials]. This may reflect
that infants were sometimes displaced or picked up by their mother during the test. However, this difference in trial completion cannot account for our overall results, as we found that while juveniles and adults exhibited habituation, but both infants and older monkeys did not.

**Baseline tendency to look**

A second possible explanation for the lifespan patterns of gaze following is that cohorts differed in their baseline tendency to look up. In particular, juveniles may have looked up more often than other cohorts in our experiment simply because they made more baseline upward looks in general. If this was the case, we predicted that juveniles would exhibit differences in their latency to look up after the experimenter did so, as well as the total duration of their looks. In terms of response latency, baseline looks should be equally likely to occur at any time during the 10s trial, whereas gaze-following responses would be more likely to occur following the experimenter’s action. Thus, if juveniles made more baseline looks than other groups, they should show longer response latencies. To address this issue, we examined trial one performance for those individuals who did follow gaze. In fact, in an ANOVA comparing performance across age cohorts, we found no significant difference in response latency in the different groups [mean response latency for infants: 3.09 ± 0.74s; juveniles: 1.88 ± 0.20s; adults: 2.13 ± 0.22s; older adults: 2.33 ± 0.55s; F$_{3,236} = 1.44$, p > 0.2, n.s.]. Indeed, if anything juveniles exhibited a slightly faster response than other groups. Furthermore, baseline looks upward are not oriented towards finding a specific target, and therefore should have a shorter duration than gaze-following responses. However, we similarly found no difference in the total duration of gazing upwards across cohorts [infant mean: 2.04 ± 0.44; juveniles: 2.14 ± 0.18; adults: 1.59 ± 0.14; older adults: 2.05 ±0.71; F$_{3,236} = 1.95$, p > 0.12, n.s]. Thus, these results do not support the possibility that differences in baseline looking rates account for our main findings.

**Study 2 Supplementary Results**

As reported in the main text, we examined how responses on *Up* and *Down trials* related to the subject's age by modeling responses for each trial type separately. For each trial type, we first fitted a base model with random subject intercepts, and condition order (first trial up or down, counterbalanced across subjects) as predictors as in the previous analyses; we did not include trial number as a covariate in this analysis as the trials alternated between the two possible types (up and down). In a second model, we then added age as an additional predictor to test its importance. We found that including age significantly improved model fit when modeling responses to *Up trials* [$\chi^2 = 14.62$, df = 1, p < 0.001]: monkey’s gaze following responses declined with age, replicating the basic results from Study 1 (see Table S5 for parameters from the full model).

<table>
<thead>
<tr>
<th>Factor</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition order (Down baseline)</td>
<td>-1.127</td>
<td>0.528</td>
<td>-2.137</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>Age (covariate)</td>
<td>-0.211</td>
<td>0.070</td>
<td>-3.000</td>
<td>&lt; 0.005</td>
</tr>
</tbody>
</table>

**Table S5:** Factors influencing propensity to follow gaze in *Up trials* (Study 2). Parameters from the full model; including age improved model fit compared to the base model.

In contrast, including age did not increase model fit for *Down trials* [$\chi^2 = 1.61$, df = 1, p > 0.20, n.s.]: baseline rates of looking upwards remained at similar low rates across ages (see Table
S6 for parameters from the full model). Consequently, the main results concerning the developmental changes in gaze-following from Study 1 are unlikely to be due to shifts in baseline reactivity.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Estimate</th>
<th>S.E.</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition order (Down baseline)</td>
<td>-0.507</td>
<td>1.404</td>
<td>-0.361</td>
<td>0.72</td>
</tr>
<tr>
<td>Age (covariate)</td>
<td>-0.238</td>
<td>0.209</td>
<td>-1.135</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Table S6: Factors influencing propensity to follow gaze in Down trials (Study 2). Parameters from the full model; including age did not improve model fit compared to the base model.

References
Supplemental Video Captions

**Video S1: Infant following gaze.** This monkey was categorized in the infant cohort (less than one years of age). The demonstrator (E1) and camera person (E2) stood next to eat other, approximately 1-2 m away from the monkey. E1 can be heard calling the monkey’s attention. When the monkey looks at her face, she says “now” and looks straight up, holding that position for the rest of the trial. E2 films the monkey’s face and times the trial, saying “stop” after a full 10s has passed.

**Video S2: Multiple discrete looks.** In this trial, the adult male monkey makes three discrete looks upwards: looking up, looking away in a different direction, and then looking up again. Multiple discrete looks was one index of the cognitive processes underlying monkey gaze following, as it reflects attempts to locate the (absent) target.