



Rhesus monkeys use both eye and head gaze to reallocate covert spatial attention facilitating visual perception

Masih Shafiei^{1,2} · Matthias Reik¹ · Marius Görner^{1,2,3,6} · Nick Taubert⁵ · Martin Giese^{4,5} · Peter Thier^{1,4}

Received: 27 August 2025 / Accepted: 26 November 2025 / Published online: 25 March 2026
© The Author(s) 2026

Abstract

Nonverbal cues, particularly eye-gaze, significantly shape human social interactions. Although nonhuman primates reliably follow head gaze, their capacity to use eye-gaze alone for inferring the other's focus of attention remains debated. We investigated this question using a realistic rhesus monkey head avatar that directed its gaze toward one of two LEDs (left or right), employing either eye movements alone or combined eye and head movements. After a randomly chosen interval (range: 50–400 ms) from gaze presentation, one LED transiently increased its luminance to near-threshold levels. Rhesus monkeys were trained to detect and report this luminance change via a saccade to the corresponding LED, independent of the avatar's gaze direction, to receive rewards. Our results showed that head-gaze cues robustly directed covert attention toward gaze-congruent targets with short delays, indicative of reflex-like, stimulus-driven orienting. In contrast, eye-gaze alone, at comparable amplitudes, did not affect attentional shifts. However, increasing the avatar's size and eye-gaze amplitude, simulating a close-range interaction, made eye-gaze cues effective in guiding attention. These findings demonstrate that rhesus monkeys possess the capacity to use eye-gaze cues to determine conspecifics' attentional targets, and validate and underscore the utility of 3D animal models as powerful tools for generating realistic yet precisely controlled stimuli. Our study supports the idea that eye-gaze following is not uniquely human but is an evolutionarily ancient ability, likely shared across Old World monkeys and apes that diverged more than 30 million years ago.

Keywords Gaze-following · Social attention · Visual spatial attention · Covert attention · Gaze perception

Masih Shafiei and Matthias Reik contributed equally to this work.

✉ Peter Thier
thier@uni-tuebingen.de

¹ Cognitive Neurology Lab, Hertie Institute for Clinical Brain Research, Eberhard Karls University of Tübingen, Otfried-Müller-Straße 27, 72076 Tübingen, Germany

² Graduate Training Centre of Neuroscience, International Max Planck Research School, Eberhard Karls University of Tübingen, Otfried-Müller-Straße 27, 72076 Tübingen, Germany

³ IMPRS for Cognitive and Systems Neuroscience, Tübingen, Germany

⁴ Werner Reichardt Centre for Integrative Neuroscience, Eberhard Karls University of Tübingen, Otfried-Müller-Straße 25, 72076 Tübingen, Germany

⁵ Section Computational Sensomotrics, Department N3, Hertie-Institute for Clinical Brain Research and Centre for Integrative Neuroscience, University Clinic Tübingen, Otfried-Müller-Straße 25, 72076 Tübingen, Germany

⁶ Present Address: Department Emotion Research, Max Planck Institute of Psychiatry, 80804 Munich, Germany

Introduction

While language serves as the primary channel of communication among humans, nonverbal cues offer a parallel and richly informative modality that leverages additional sensory and motor systems for both signal transmission and reception (Mehrabian, 2017). In verbal communication, the sender primarily engages articulatory mechanisms, while the receiver relies on the auditory system. In contrast, nonverbal communication recruits the sender's bodily movements, such as facial expressions, gestures, and posture, and mainly the receiver's visual system to convey and interpret meaning. This multimodal interaction enables the receiver to integrate verbal and nonverbal information, thereby enhancing the efficiency, reliability, and perhaps depth of communication (Drijvers & Holler, 2023). Nonverbal cues also provide a means to assess the congruence and trustworthiness of verbal content, allowing the receiver to cross-reference messages across modalities (Morioka et al., 2016). Importantly, nonverbal cues are not merely supplementary to language

but can operate independently to convey information. For example, an individual can silently draw another's attention to an object of interest by first establishing eye contact and then shifting their gaze toward the target. This sequence prompts the observer to follow the initiator's gaze and locate the object, illustrating how attention can be directed toward targets through nonverbal cues alone (Bock et al., 2008). Among various nonverbal cues, gaze direction, defined by the spatial orientation of the other's eyes, face, or body, serves as a salient indicator of their current focus of attention. The ability to align one's own attention with the other's gaze, known as gaze following, allows individuals to access relevant information in the environment, such as potential threats or resources that might otherwise go unnoticed (Butterworth & Jarrett, 1991). Moreover, gaze following supports higher-order social cognitive processes by enabling individuals to infer others' mental states, such as intentions, desires, or beliefs, a foundational capacity that helps humans form a Theory of Mind (ToM; Charman et al., 2000; Sodian & Kristen-Antonow, 2015). In humans, gaze following has been strongly associated with first language acquisition (Beuker et al., 2013; Brooks & Meltzoff, 2008; Morales et al., 1998), and impairments in this ability have been linked to delays in social development, particularly among individuals with autism spectrum disorder (ASD; Congiu et al., 2016; Dawson et al., 2004; Elsabbagh et al., 2012; Thorup et al., 2016).

Previous research has explored the evolutionary origins of gaze-following by studying a range of animal species, whose evolutionary trajectories diverged from our own species at different times. Nonhuman primates (NHPs) have been of particular interest given their close evolutionary relationship with humans and their ability to engage in complex social behaviors (Cheney et al., 1986; Platt et al., 2016). The goal has been to examine the extent of similarity between the gaze-following abilities of humans and NHPs. Although it has been consistently shown that both humans and NHPs follow the gaze direction of either conspecifics or humans to co-orient their attentional focus with the other's, the type of gaze cues used varies across species (Anderson & Mitchell, 1999; Bräuer et al., 2005; Burkart & Heschl, 2006; Byrnit, 2009; Deaner & Platt, 2003; Emery et al., 1997; Ferrari et al., 2000, 2008; Goossens et al., 2008; Itakura, 1996; Itakura & Tanaka, 1998; Kano & Call, 2014; Kano et al., 2022; Kaplan & Rogers, 2002; Liebal & Kaminiski, 2012; Marciniak et al., 2015; Parron & Meguerditchian, 2016; Peignot & Anderson, 1999; Povinelli & Eddy, 1996, 1997; Rosati et al., 2016; Ruiz et al., 2009; Santos & Hauser, 1999; Shepherd & Platt, 2008; Teufel et al., 2010; Tomasello et al., 1998, 1999, 2007; Yu et al., 2012). Humans are capable of accurately following eye-gaze direction in 3D space by exploiting eye direction relative to the head together with information on head and

body orientation (Atabaki et al., 2015; Bock et al., 2008; Moors et al., 2015). This ability has long been linked to the distinctive morphology of the human eye characterized by a very bright sclera and a horizontally elongated shape that exposes more of the sclera, features thought to make the eye markedly more conspicuous than that of NHPs (Kobayashi & Kohshima, 1997, 2001). However, recent quantitative work contests this view; when eye visibility is studied with finer metrics such as iris-to-sclera luminance contrast or the proportion of sclera revealed during averted (rather than direct) gaze, human eyes are no longer significantly different from several NHP species (Mayhew & Gómez, 2015; Mearing & Koops, 2021; Perea-García et al., 2019, 2025). On the other hand, although previous reports consistently support NHPs' ability to use the other's head direction to identify their object of interest (Anderson & Mitchell, 1999; Bräuer et al., 2005; Burkart & Heschl, 2006; Byrnit, 2009; Deaner & Platt, 2003; Emery et al., 1997; Ferrari et al., 2000; Goossens et al., 2008; Itakura, 1996; Itakura & Tanaka, 1998; Kano & Call, 2014; Liebal & Kaminiski, 2012; Marciniak et al., 2015; Parron & Meguerditchian, 2016; Peignot & Anderson, 1999; Povinelli & Eddy, 1996, 1997; Rosati et al., 2016; Ruiz et al., 2009; Shepherd & Platt, 2008; Teufel et al., 2010; Tomasello et al., 1998, 1999, 2007), there is still no clear consensus on whether they can follow the others eye-gaze in the absence of accompanying head turns (Burkart & Heschl, 2006; Deaner & Platt, 2003; Ferrari et al., 2000; Itakura & Tanaka, 1998; Kano et al., 2022; Lorincz et al., 2001; Peignot & Anderson, 1999; Povinelli & Eddy, 1996, 1997; Santos & Hauser, 1999; Tomasello et al., 2007; Yu et al., 2012). This lack of agreement may stem from methodological differences across studies. First, several studies have relied on gaze cues from live human demonstrators (Burkart & Heschl, 2006; Ferrari et al., 2000; Itakura & Tanaka, 1998; Peignot & Anderson, 1999; Povinelli & Eddy, 1996, 1997; Santos & Hauser, 1999; Tomasello et al., 2007), which are likely unnatural for NHPs. In others, static images of conspecifics were used (Deaner & Platt, 2003; Kano et al., 2022; Lorincz et al., 2001) instead, but such stimuli lack the dynamic motion inherent to natural gaze cues, a feature that has been demonstrated to influence gaze-following behavior (Anderson et al., 2016). Second, some studies examined gaze-following responses in non-object-oriented tasks (Ferrari et al., 2000; Kano et al., 2022; Lorincz et al., 2001; Povinelli & Eddy, 1997; Tomasello et al., 2007), known to attenuate orienting responses with repeated exposure (Parron & Meguerditchian, 2016; Tomasello et al., 2001), or linked gaze direction to object detection in a reward-based task (Burkart & Heschl, 2006; Byrnit, 2009; Itakura & Tanaka, 1998; Peignot & Anderson, 1999), potentially training

animals to associate gaze with reward rather than revealing spontaneous or intrinsic gaze-following abilities. Finally, a key question is whether eye direction relative to the head is used by NHPs as a vectorial cue that indicates spatial direction, or merely as a modulatory signal that enhances or interferes with the processing of head-derived information. This distinction is essential when considering whether NHPs' limited use of eye-gaze cues reflects a perceptual constraint in processing eye-gaze or simply poorer input quality, especially when the visibility of the eyes may be reduced. If the latter explanation holds, eye-gaze may play a more significant role where its informational value is highest such as when social interactions occur at a closer range.

To address these limitations of the previous research, we designed the current study. To overcome the challenge of providing ecologically valid yet controlled gaze cues, we decided to deploy a hyper-realistic rhesus monkey head avatar as demonstrator. This avatar, known to elicit completely natural behavioral reactions in monkey observers (Siebert et al., 2020), allowed precise control over gaze direction, as well as control over the relative contributions of head and eye orientation, all while maintaining natural motion dynamics. The use of the avatar enabled us to investigate our core question: whether eyes alone can modulate the observer's focus of spatial attention. To this end, we used variants of gaze shifts presented by the avatar: a first one in which the avatar turned its head towards a target with the eyes moving with the head and a second one in which only the eyes were directed to the target whereas the head was kept straight. To address the second limitation of previous work, we designed a paradigm that ensured gaze-following itself was not reinforced, isolating the possible influence of spontaneous gaze-following behavior on attention and perception. Specifically, we implemented a near-threshold luminance detection task, preceded by a gaze cue. The avatar randomly directed the gaze to one of the two LEDs, placed to the right and left. After a variable time from the gaze cue, one of the LEDs randomly increased its luminance to a near-threshold level for a brief duration. The rhesus monkeys were trained to detect the change and report it, regardless of the avatar's gaze direction, by making a saccade to the corresponding LED to receive reward for correct responses. The assumption was that the gaze cue induced shifts of covert attention, as the animals were required to maintain fixation on the avatar, which should manifest themselves in an improvement of luminance change detection thresholds. The covert shift of attention is the suggested underlying mechanism as the animals were prohibited from making overt attention shifts—based on foveating saccades—to the location cued by the gaze stimuli. In other words, the improvement in performance occurred because the near-threshold luminance cue coincided with the new location of attention. Importantly, our use of the term *covert attention* refers specifically to nonsaccadic shifts of visual attention not discernible by an observer as defined in the visual attention literature (e.g., Carrasco, 2011;

Posner, 1980), and should not be confused with its use in the NHP eye-gazing literature, where the term has occasionally been associated with the hypothesis that scleral pigmentation could conceal gaze direction. We examined a range of stimulus onset asynchronies (SOAs), spanning from very short (50 ms) to relatively long (400 ms). The reason was that previous research had shown that exogenous, stimulus-driven shifts of attention occur rapidly and are most evident at shorter SOAs, while endogenous, goal-directed modulation of attention develops more slowly and is therefore apparent at longer SOAs. Accordingly, our inclusion of multiple SOA conditions was intended to probe both types of attentional modulation in response to the gaze stimulus. We show that head gaze cues draw the observer's focus of attention to objects of interest with short delays which are fully compatible with reflex-like stimulus-driven behavior, but not so for eye-gaze shifts of the same amplitude in the absence of accompanying head movements. However, also eye-gaze shifts become effective if the distance to the avatar head is substantially reduced and the amplitude of the eye-gaze increased, simulating close-range social interactions with improved eye visibility. This manipulation was implemented by increasing the amplitude of the eye-only gaze presented with a proportionally larger avatar size. We note that this condition was tested in only one monkey due to practical constraints (described in detail in the Methods) and should therefore be treated as an exploratory result.

Methods

Experimental subjects

The selection of rhesus monkeys (*Macaca mulatta*) as the animal model for this study was guided by our aim to investigate the evolutionary origins of human gaze-following by exploring the gaze-following capabilities of rhesus monkeys, arguably the best-explored representative of old world monkeys whose visual and visuomotor systems match those of humans to a large extent (Lu et al., 2023). We used two male monkeys, referred to as “Monkey E” and “Monkey C.” Monkey E was naïve to tasks resembling those used in the present study. Monkey C had previously undergone partial training for an unrelated project involving static images of conspecific head-gaze cues, in which he was required to either follow or ignore the gaze cue based on a rule dictated by a precue. However, his training was not completed due to unrelated circumstances. Both animals had pre-existing headpost implants, which enabled the painless fixation of the head necessary for high-precision eye-tracking (for a description of the detailed neurosurgical procedures, see Prsa & Thier, 2011). Additionally, Monkey E had received a scleral search coil implant, allowing eye tracking using an in-house system developed based on a design originally

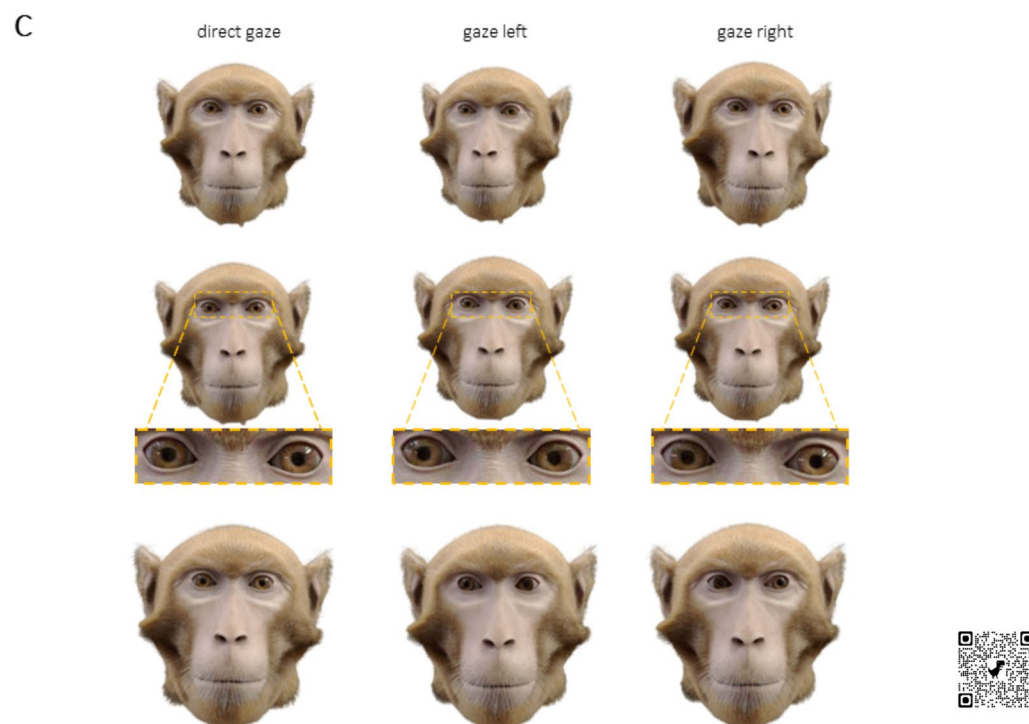
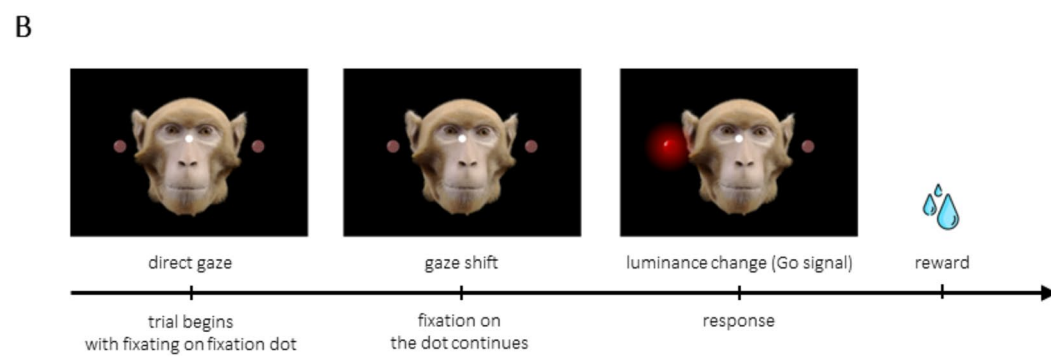
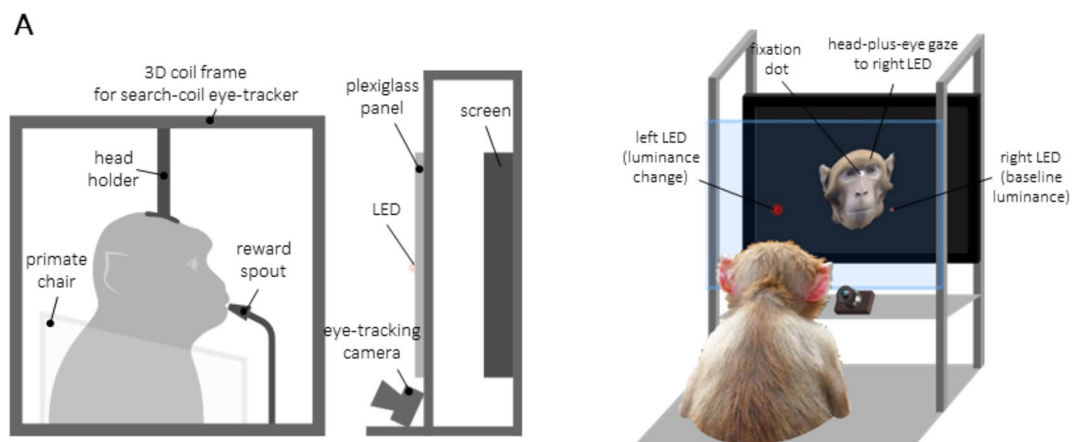


Fig. 1 **A** Schematic of the experimental setup. The coil frame used for search-coil eye tracking, shown on the left panel, was employed exclusively for Monkey E. On the right, a simplified illustration of the experimental setup is shown; components such as the head-fixation apparatus, reward delivery system, and primate chair have been omitted for clarity. **B** Sequence of events in a trial. This panel shows an incongruent trial, where the avatar's gaze is directed towards the distractor LED. **C** Each row represents a different gaze-cue type: head gaze (top), small-amplitude eye-gaze (middle), and large-amplitude eye-gaze (bottom). Columns show video snapshots depicting, from left to right: a direct gaze, a gaze toward the left LED, and a gaze toward the right LED. Insets in the middle row provide a zoomed-in view of the peri-orbital region to enhance visualization of the gaze direction. Videos of the different gaze cues are accessible via the QR code in panel C, or through the following link: <https://doi.org/10.6084/m9.figshare.28769981> (Color figure online)

described by Bechert and Koenig (1996). Eye movements of Monkey C were recorded via high-resolution video-based eye tracking using the EyeLink 1000 Plus desktop system (SR Research Ltd., Canada).

On experimental days, animals were maintained under controlled water-access protocols, which ensured that water was received as a positive reinforcement upon successful trial completion. To enhance motivation, Monkey C was occasionally rewarded with apple juice instead of water because, unlike Monkey E, he was very fond of juice.

Experimental setup

Figure 1A illustrates a schematic of the experimental setup. Head-fixed animals, calmly seated in a primate chair, were positioned inside a dimly lit, sound-attenuated booth. A 24-inch LCD monitor (BenQ XL2411-B; 1,920 × 1,080 resolution; 60 Hz refresh rate) was mounted 60 cm in front of the animals, centered on their neutral line of sight. A transparent plexiglass panel was placed 14 cm in front of the screen. Two LEDs were fixed on this panel at 7.5 cm eccentricity to the left and right of the center along the horizontal axis. The LEDs were illuminated at a baseline luminance of 0.09 cd/m² for Monkey E and 0.01 cd/m² for Monkey C. Rewards were delivered directly to the animal's mouth via a spout positioned in front of the animal, controlled by a fully automated, in-house-developed, computer-controlled reward system.

Behavioral paradigm

Figure 1 illustrates the trial structure and the different gaze-cue conditions. We generated gaze cues using a monkey head avatar previously validated by Siebert et al. (2020). The 3D head model was constructed from a structural MRI scan of a particular male rhesus monkey, with additional refinements—such as skin texture, fur, the coloration of different facial regions and the features of the eyes—based on

reference photographs taken from our own rhesus monkeys (for details, see Siebert et al., 2020).

At the start of each experimental session, eye movements were calibrated using a standard nine-target calibration paradigm. The LEDs were kept off during calibration. Each trial in the main behavioral paradigm, designed to elicit gaze following, began with the appearance of a forward-facing avatar, with a white fixation dot (0.2° diameter; CF = central fixation dot) on its nose bridge as shown in Fig. 1A and B. This dot was horizontally aligned midway between the two LEDs. To initiate a trial, monkeys were required to maintain fixation within an invisible 5° × 5° window around the CF for 500 ms.

Following successful fixation, the avatar randomly shifted its gaze toward one of the two LEDs. After a variable delay, referred to as the stimulus onset asynchrony (SOA), the luminance of one of the LEDs increased briefly (for 300 ms), thereby identifying the target LED. The target LED was chosen at random, independently of the avatar's gaze direction. Its luminance rose to 0.28 cd/m² (from 0.09 cd/m²) for Monkey E and 0.07 cd/m² (from 0.01 cd/m²) for Monkey C yielding Δ Lum (delta luminance) values of 0.18 and 0.06, respectively. These individualized Δ Lum values were determined for each animal based on a perceptual threshold estimation procedure that identified the luminance change required to achieve 75% detection accuracy (see Methods section 'Perceptual Threshold Estimation'). The onset of the Δ Lum stimulus defined the earliest time point the monkey could report the target LED by making a saccade toward it, regardless of the avatar's gaze direction. Correct responses were rewarded, whereas incorrect choices (saccades made to the distractor LED) were not. After either a correct or incorrect response, an intertrial interval (ITI) ranging between 700–1,500 ms was introduced, during which the avatar's gaze returned to its neutral position. Trials were self-paced allowing monkeys to initiate the next trial by fixating on the CF.

If a monkey broke fixation after the gaze cue but before Δ Lum, the same trial was repeated. Trials were aborted if the monkey failed to respond within 1.5 s of the avatar's gaze shift. To help Monkey C differentiate between premature responses and incorrect choices, an aversive buzz tone was played when the monkey responded too early, before the Δ Lum appeared. Monkey E, whose data collection was completed earlier and who showed no difficulty in making this distinction, did not require the use of the buzz tone (see the "Animal Training" section for more details).

To examine the temporal dynamics of attention, we tested five different SOAs: 50, 100, 200, 300, and 400 ms. In each session, a subset of two or three SOAs was selected pseudorandomly, and one SOA from the subset was randomly assigned to each trial. Three distinct gaze cue types were tested (directed toward either left or right):

eye-gaze accompanied by head turns (Fig. 1C, top row), eye-gaze without head movement (Fig. 1C, middle row), large-amplitude eye-gaze without head movement, featuring a larger avatar head (Fig. 1C, bottom row). Only one gaze cue type was presented per session. Videos of the gaze stimuli can be accessed via the QR code or URL provided in Fig. 1. In Monkey C, under the eye-only gaze condition with small amplitude, we tested three SOAs (50, 200, and 400 ms), chosen to cover the range most likely to produce significant effects based on previous findings on head-gaze following. We intended to test the remaining two SOAs (100 and 300 ms) in case the results obtained would require further refinement, however this proved unnecessary.

Initially, we compared responses to the avatar producing a 7.7° gaze shift by either moving eyes and head together with the eyes kept straight relative to the head (head-gaze; Fig. 2c, top row) or only the eyes (eye-gaze; Fig. 2c, middle row) while the head stayed put (Fig. 1C). Next, we implemented a more salient eye-gaze cue by enlarging the avatar by 40%, while keeping all other experimental parameters constant. For the identical gaze angle of the avatar, this manipulation entailed a larger translation of the avatar pupil on the screen from the straight-ahead video frame to the postsaccadic video frame. For the observer, this led to a 40% increase in the perceived gaze-shift angle (10.8° to either side). While Monkey E was able to cope with this configuration, Monkey C frequently broke fixation even before the gaze cue was presented, forcing us not to consider this monkey any further in this experiment. We hypothesize that the larger avatar head in this condition may have been experienced as overly dominant and threatening, overwhelming the relatively small-sized Monkey C. In support of this interpretation, we observed that Monkey C immediately resumed task engagement and fixated successfully upon switching the avatar back to its original size in which the head appeared smaller again. Note that we did not attempt to collect responses to larger eye-gaze shifts for this latter configuration as it would have been required to move the LEDs further out on the transparent panel.

Animal training

The monkeys were initially trained to associate an increase in LED luminance from the fixed baseline value with a reward. This luminance change (ΔLum) was randomly presented by one of two LEDs. Task difficulty was gradually increased by reducing both the magnitude and duration of the ΔLum . Once the monkeys consistently associated brief changes in LED luminance, regardless of absolute luminance levels or LED position, with a reward, they progressed to the next phase of the experiment: perceptual threshold assessment.

Perceptual threshold assessment

This phase of the study aimed to determine the ΔLum corresponding to each animal's point of subjective equality (PSE), defined as the luminance change that produced an average 25% increase in correct responses, relative to the baseline chance level of 50%. To achieve this, we systematically varied ΔLum , while the duration of luminance change was held constant at 300 ms, enabling an assessment of the relationship between ΔLum and detection accuracy. A simplified version of the main behavioral paradigm was employed during this phase. To simulate the light emitted by the avatar in the main experiment, a colored silhouette of the avatar, matched in luminance, was displayed throughout the perceptual threshold assessment trials (see Fig. S1A).

Each trial began with the appearance of a CF superimposed on the silhouette. The monkey was required to fixate on the CF for the luminance change event to occur. In Monkey E, the luminance changes of the LED occurred after a variable delay, randomly chosen from 50, 100, 200 ms, following the CF presentation. In contrast, Monkey C experienced a fixed temporal delay of 300 ms between the CF presentation and the luminance change. In both monkeys, the luminance change lasted 300 ms and appeared randomly at one of two possible LEDs. Following the luminance change, the monkeys had a maximum of 1,500 ms to initiate a saccadic response. The ΔLum values ranged from high levels, yielding near-perfect detection accuracy, to low levels, resulting in a detection accuracy no better than chance. Correct responses were rewarded with immediate positive reinforcement (a drop of water). An ITI of 700–1,500 ms separated trials.

The ΔLum values corresponding to the PSE were 0.18 cd/m^2 for Monkey E and 0.06 cd/m^2 for Monkey C, which were subsequently used in the main experiment (see Fig. S1B). The observed differences in PSE between subjects likely reflect individual variations in their sensitivity to perceive changes in luminance (e.g., due to noncorrected ametropia, mild enough to not interfere with sufficiently precise fixation and saccades).

Data processing and statistical analysis

Data preprocessing was conducted using a custom-written script in Julia (Version 1.10.3). Statistical data analysis and visualization were performed in R (version 4.4.2, 2024–10–31; R Core Team, 2024) using the following packages: *ggplot2* (version 3.5.1; Wickham, 2016), *lme4* (version 1.1.36; Bates et al., 2015), *car* (version, 3.1.3; Fox & Weisberg, 2019), *emmeans* (version 1.10.7; Lenth, 2025), and *performance* (version 0.13.0; Lüdtke et al., 2021). Estimation of ΔLum values, yielding an average detection accuracy of 75% was carried out in MATLAB (R2023a; The

MathWorks, Natick, MA, USA) using the *Psignifit* toolbox (version 4). Hit rates for each Δ Lum level were modelled using a logistic psychometric function, with the lower asymptote fixed at 50% to reflect chance-level performance.

Eye-tracking data, recorded as a time series of horizontal eye positions, were smoothed using a Savitzky-Golay filter (11 ms window, first-order polynomial). The first derivative of the smoothed eye position signal was computed to extract the onset and direction of the response saccade. Reaction time (RT) was defined as the temporal interval between Δ Lum onset and saccade onset. Trials with reaction times below 70 ms or above 1,500 ms were excluded, as they were considered unrelated to the task. This resulted in the removal of 5.7% of trials for Monkey E (5.3% in the head-gaze condition, 5.9% in the small-amplitude eye-gaze, and 2.6% in the large-amplitude eye-gaze) and 10.7% for Monkey C (11.7% in the head-gaze, and 9.7% in the small-amplitude eye-gaze). The final dataset comprised 9,717 head-gaze trials, 35,830 small-amplitude eye-gaze trials, and 26,885 large-amplitude eye-gaze trials for Monkey E. For Monkey C, 10,154 head-gaze trials and 10,820 small-amplitude eye-gaze trials were retained for analysis.

To assess the impact of gaze cue congruency on detection accuracy, generalized linear mixed models (GLMMs) with a logistic link function and maximum-likelihood estimation were used. The dependent variable was *hit*, coded as a binary outcome (1 = correct, 0 = incorrect). Fixed predictors included *congruency* (congruent vs. incongruent), SOA (50, 100, 200, 300, and 400 ms), target direction (left vs. right), normalized log-transformed RT, SOA $\times$ congruency, SOA $\times$ target direction, and SOA $\times$ normalized log-transformed RT. Reaction times were log-transformed and *z*-scored prior to analysis. For Monkey C, only three SOAs (50, 200, and 400 ms) were used in the eye-only condition as noted earlier. Target direction was specifically included to detect potential lateralized biases in performance. The GLMM estimates coefficients for fixed predictors in log-odds units, representing the log odds ratio (*OR*) of a correct detection given a one-unit change in a continuous predictor or a comparison between a given and reference level for categorical predictors. Where appropriate, log-*OR* values were converted to odds ratios (*ORs*) using the exponential function to facilitate interpretation.

Model selection followed a backward stepwise elimination approach, using data from the first subject (Monkey E) as an exploratory dataset. At each iteration, predictors were removed if their *p*-values were ≥ 0.05 and/or, for predictors with one degree of freedom (*df*), if their generalized variance inflation factor (GVIF) exceeded 3.2 (Fox & Monette, 1992; Fox & Weisberg, 2019). After each elimination step, the model was re-estimated and re-evaluated using the same criteria. This process generated a series of candidate models, which were compared using the Akaike information criterion

(AIC; Symonds & Moussalli, 2011). The model with the lowest AIC, provided that all retained predictors with *df* = 1 had GVIF < 3.2, was selected as the best-fitting model. A Δ AIC > 10 was considered strong evidence in favor of the model with the lower AIC (Burnham et al., 2004). The significance of the random effect was assessed by comparing the final GLMM to an equivalent linear model (LM) without the random term, based on AIC. In addition, the 95% confidence interval (CI) for the variance explained by the random effect was computed; the random effect was retained only if this CI did not include zero (Bates et al., 2015; Nakagawa & Schielzeth, 2013). We also report the adjusted intraclass correlation coefficient (ICC) for the random factor in the best-fitting GLMMs to quantify the proportion of residual variance attributed to session-level effects after accounting for fixed predictors. The final models derived from Monkey E were then applied to data from Monkey C to assess generalizability. Consistency in the direction and statistical significance of key predictors across both subjects was interpreted as evidence of model robustness and replicability.

To examine the influence of congruency on hit rates within each SOA level, post hoc pairwise comparisons between congruent and incongruent conditions were performed using the estimated marginal means (EMMs) derived from the best-fitting models. *P*-values from these comparisons were corrected for multiple testing using the Benjamini–Hochberg false discovery rate (FDR) procedure (Benjamini & Hochberg, 1995). Additionally, to evaluate the overall contribution of each fixed effect in the final models, Type III Wald chi-square tests were conducted using analysis of deviance (Type III ANOVA; Fox, 2015; Luke, 2017). *P*-values from these tests were also corrected for multiple comparisons using the Benjamini–Hochberg method. We report the Wald χ^2 statistic for each fixed effect, which quantifies the extent to which the inclusion of a specific predictor improves model fit, conditional on all other variables in the model.

Results

Effect of head-gaze on detection accuracy

We found a significant main effect of congruency, SOA, and target direction along with significant interactions between SOA and target direction, as well as between SOA and RT, on hit probability in both monkeys (Table 1 and Fig. 2). Post hoc pairwise comparisons revealed a significantly higher hit probability in the congruent condition compared with the incongruent condition at SOAs of 50, 100, and 200 ms in both monkeys (see table 1 and Fig. 2 and 3). This finding suggests that target detection is significantly enhanced when the target is congruent with the preceding head-gaze cue,

Table 1 GLMM and ANOVA outputs for the fixed predictors in the eye-head gaze condition

	Monkey E ^a		Monkey C ^b	
	Wald χ^2	log-odds	Wald χ^2	log-odds
Intercept	206.11 ***	1.62 ***	144.61***	1.26 ***
Congruency (ref: incongruent)	12.34 ***	0.34 ***	4.56*	0.22*
SOA (ref: 50)	18.89 ***	-	12.32*	-
100	-	-0.16	-	0.14
200	-	-0.74 ***	-	-0.14
300	-	-0.23	-	0.28
400	-	-0.56 *	-	0.36
Target direction (ref: right)	11.17 ***	-0.32 ***	5.04*	-0.23*
Normalized log RT	1.59	0.08	6.16*	0.20*
SOA $\times$ congruency	3.65	-	20.36***	-
100:congruency	-	0.07	-	0.07
200:congruency	-	0.19	-	0.11
300:congruency	-	-0.16	-	-0.50**
400:congruency	-	-0.03	-	-0.35*
SOA $\times$ target direction	74.64 ***	-	115.79***	-
100:target direction	-	-0.16	-	-0.14
200:target direction	-	-0.32	-	-0.67 ***
300:target direction	-	-1.18 ***	-	-1.25 ***
400:target direction	-	-0.88 ***	-	-1.22 ***
SOA $\times$ normalized log RT	15.36 **	-	27.76***	-
100: normalized log RT	-	0.03	-	-0.24*
200: normalized log RT	-	0.26 *	-	0.23*
300: normalized log RT	-	0.20 *	-	0.22*
400: normalized log RT	-	0.3 **	-	0.26**
Congruency $\times$ normalized log RT	0.46	-0.04	5.22*	-0.13*

^aNumber of observations: 9245; random predictor (session number = 25): variance = 0.107; adjusted ICC = 0.032

^bNumber of observations: 9518; random predictor (session number = 16): variance = 0.027; adjusted ICC = 0.011

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

provided that the cue–target interval does not exceed 200 ms. The odds ratios (ORs) indicated that the likelihood of a hit in the congruent condition was 1.4–1.7 times higher in Monkey E and 1.25–1.4 times higher in Monkey C, depending on the SOA (see Table 1 and Fig. 2). Furthermore, at an SOA of 300 ms, Monkey C exhibited a significant inverse congruency effect, with hit probability being 1.14 times higher in the incongruent condition than in the congruent condition, suggesting a potential inhibition-of-return (IOR) effect. IOR refers to the temporary inhibition of attentional reorientation toward a recently cued location (Gazzaniga et al., 2019).

Regarding reaction time, the analysis revealed that faster RTs were associated with lower hit probability: a one-standard deviation decrease in normalized, log-transformed RT (≈ 624 ms faster) reduced the odds of a correct response by 23.3% overall (22.8% in Monkey E, and 23.7% in Monkey C). This effect was moderated by congruency and SOA. Given the temporal uncertainty introduced by variable SOAs

and the low salience of the luminance transient (reduced intensity and brief presentation), hit probability increased with longer response latencies, likely allowing additional processing before the response was made. However, in congruent trials, the reduction in hit probability associated with faster RTs was significantly attenuated; for Monkey C, the decrease in hit probability was reduced by approximately 14% at shorter SOAs (≤ 200 ms). Monkey E showed a similar pattern, with a nonsignificant reduction of about 25% across all SOAs. Thus, congruency mitigated—but not eliminated—the negative association between RT and hit probability probability. In congruent trials, the gaze cue provided supplementary spatial information that supported target detection. The covert shift of attention triggered by the gaze direction increased the probability that the brief luminance change occurs within the attentional spotlight. Consequently, faster RTs in these trials are likely affected by facilitated detection of the luminance change besides the

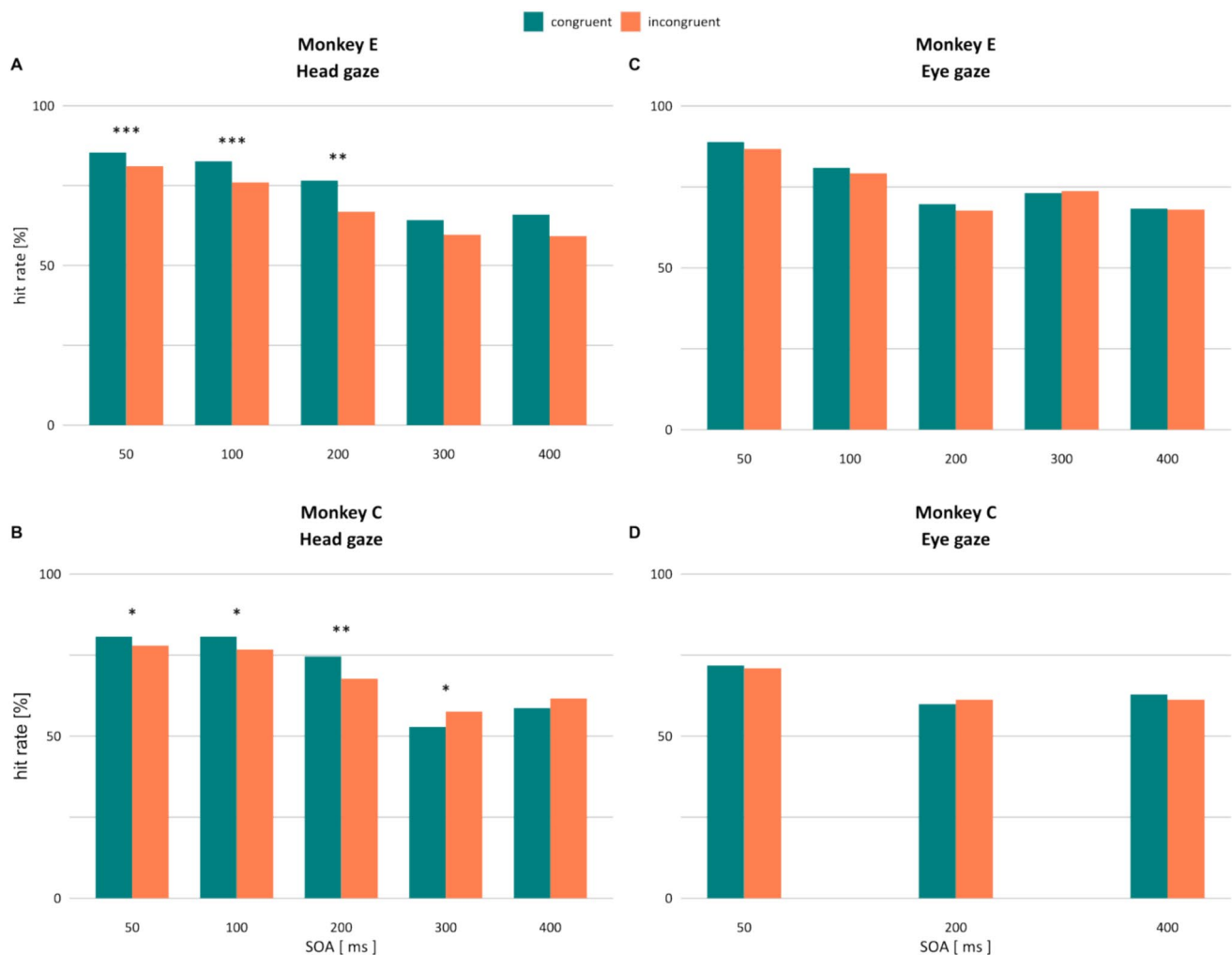


Fig. 2 Hit rate as a function of SOA and congruency. Panels **A** and **B** show results from the eye-head condition, while panels **C** and **D** display results from the small-amplitude eye-only condition. Data for Monkey E are presented in the top row (**A**, **C**), and data for Monkey

C in the bottom row (**B**, **D**). A significant effect of congruency on hit rate is observed in the eye-head condition but not in the eye-only condition. This pattern is consistent across both monkeys. * $p < .05$; ** $p < .01$; *** $p < .001$. (Color figure online)

random response variability, thereby altering the relationship between RT and hit probability.

The models further revealed that hit probability was systematically influenced by a target direction bias (Table 1). Specifically, the probability of a hit was significantly lower for leftward targets than for rightward targets. This effect varied across SOAs due to significant interactions between target direction and SOA. In both monkeys, the most pronounced direction bias was observed at an SOA of 50 ms ($OR = 0.73$ in Monkey E; $OR = 0.80$ in Monkey C), while the weakest bias occurred at SOA 300 ms ($OR = 0.22$ in Monkey E; $OR = 0.23$ in Monkey C). The consistent rightward bias across both monkeys, together with the comparable range of effect sizes, suggests a shared underlying mechanism influencing response patterns. To ensure that the observed effect of congruency on hit rate was not confounded by target

direction bias, we compared the effect size of congruency between the best-fitting GLMM and a reduced model that excluded target direction as a predictor. Both models exhibited identical fit statistics (AICs = 9,372; BICs = 9,528.9). Importantly, the fixed-effect estimate for congruency remained virtually unchanged between the full and reduced models ($\beta = 0.3426662$ vs. 0.3426667), corresponding to a difference in odds ratio of less than 2×10^{-7} . This confirms that target direction bias did not meaningfully influence the relationship between congruency and hit rate.

Effect of small amplitude eye-gaze on detection accuracy

We found significant main effects of congruency and SOA on hit probability in Monkey E under the small-amplitude

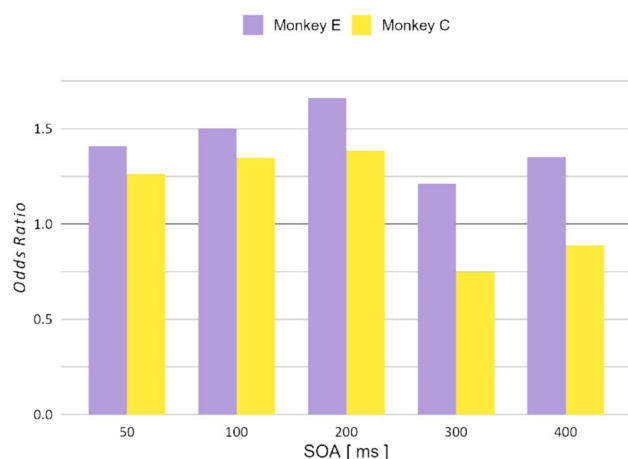


Fig. 3 Odds ratios for the effect of congruency in the head-gaze condition on hit probability as a function of SOA, shown separately for each monkey. In both monkeys, the probability of a hit in the congruent condition increases with SOA up to 200 ms. However, a notable decline in the odds ratio is observed for longer SOAs. This decline is more pronounced in Monkey C, in which the odds ratio drops below 1.0 at 300 ms, consistent with an inhibition-of-return effect. (Color figure online)

eye-only condition. In contrast, no significant main effects were observed in Monkey C (Table S1). In both monkeys, significant interaction effects were detected between SOA and target direction, as well as between SOA and normalized log-transformed RT. However, post hoc comparisons did not reveal any significant differences in hit probability between the congruent and incongruent conditions across SOAs in either monkey (Table 2 and Fig. 2).

Effect of large amplitude eye-gaze on detection accuracy

So far, our findings suggest that the type of gaze cue offered by the others (head versus eyes only) determines

whether visual spatial attention of monkeys shifted to the target. Specifically, it appears that monkeys are able to use spatial information provided by a conspecific's head-gaze but not by their eye-gaze. However, is this conclusion fully justified? Rather than reflecting a categorical difference, the inefficacy of eye-gaze might be secondary to poorer visibility of the small-amplitude eye-gaze. After all, eyes cover a much smaller area of the visual field of the observers' eyes than the face or head. This difference in size might affect how salient eyes are and therefore entail different potentials to guide attention. To assess whether the absence of an eye-gaze-dependent boost in luminance change detection might be a consequence of the low salience of the eye-gaze cue rather than a perceptual constraint in following eye-gaze, we introduced the large-amplitude eye-gaze cue condition. This involved increasing the avatar's size and the angular deviation of its gaze. All other aspects of the paradigm and experimental setup were held constant to isolate the effect of the modified eye-gaze cue on attention modulation.

Using the large-amplitude eye-gaze cue, we observed significant main effects of congruency, SOA, target direction, as well as significant interactions between SOA and target direction and between SOA and normalized log-transformed RT on the probability of a hit (Table 3 and Fig. 4). Post hoc pairwise comparisons revealed significantly higher hit probabilities in the congruent condition at SOAs of 50 ms ($OR = 1.17$, adj. $p = .04$), 100 ms ($OR = 1.18$, adj. $p = .04$), and 400 ms ($OR = 1.30$, adj. $p = .002$). These findings strongly suggest that the absence of a congruency effect on hit probability observed with the original eye-gaze cue can be attributed to the lack of conspicuity of the gaze signal. By increasing the amplitude of the eye-gaze cue, we effectively triggered gaze-following behavior, which in turn biased hit probability toward the congruent condition.

Table 2 Odds ratios (OR) and corresponding p -values for post hoc comparisons between the congruent and incongruent conditions

SOA	Monkey E				Monkey C			
	Eye-head		Eye only		Eye-head		Eye only	
	OR	p_{adj}	OR	p_{adj}	OR	p_{adj}	OR	p_{adj}
50	1.41	.001	1.22	.072	1.25	.041	1.04	.607
100	1.52	<.001	1.13	.072	1.34	.022	–	–
200	1.70	<.001	1.09	.124	1.39	.006	0.91	.493
300	1.20	.151	1.04	.659	0.75	.022	–	–
400	1.36	.151	0.99	.793	0.88	.291	1.09	.793

Odds ratios (OR) greater than 1 indicate a higher probability of a hit in the congruent condition than in the incongruent condition. Boldface values highlight significant OR s for improved readability. p_{adj} : adjusted p values

Table 3 GLMM and ANOVA outputs for the fixed predictors in the eyes-only gaze condition with large amplitude

	Monkey E	
	Wald χ^2	log-odds
Intercept	30.46 ***	0.58 ***
Congruency (ref: incongruent)	5.00 ***	−0.16 *
SOA (ref: 50)	72.41 ***	-
100	-	−0.19
200	-	0.84 ***
300	-	−0.10
400	-	0.93 ***
Target direction (ref: right)	250.29 ***	0.29 ***
Normalized log RT	0.01	0.01
SOA × congruency	10.93	-
100:congruency	-	−0.01
200:congruency	-	0.22 *
300:congruency	-	0.05
400:congruency	-	−0.07
SOA × target direction	588.86 ***	-
100:target direction	-	−0.02
200:target direction	-	−0.37 ***
300:target direction	-	−0.22 ***
400:target direction	-	−0.47 ***
SOA × normalized log RT	31.49 **	-
100: normalized log RT	-	0.16 *
200: normalized log RT	-	0.25 ***
300: normalized log RT	-	0.25 ***
400: normalized log RT	-	0.29 ***
Congruency × normalized log RT	0.28	0.02

Number of observations: 25926; random predictor (session number = 28): variance = 0.070; adjusted ICC = 0.020

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Discussion

While previous work has firmly established that rhesus monkeys use conspecifics' head-gaze to shift their focus of attention—either overtly or covertly—toward the other's object of interest, the role of the others' eye-gaze has remained contentious. Hence, in this experiment, we aimed to address whether object-oriented gaze cues demonstrated by a conspecific's eyes in the absence of accompanying head turns can guide covert shifts in visual spatial attention in rhesus monkeys. Rather than relying on videos or static images of real conspecifics performing gaze shifts, we used a realistic rhesus monkey head avatar to display object-directed eye-gaze, presented either with or without accompanying head turns. The monkey observers were then tasked with detecting near-threshold luminance changes at one of two possible visual targets, with the demonstrator's gaze directed toward

the target in 50% of trials. Our results revealed significantly higher accuracy when the demonstrator's gaze was directed toward the target object, but this effect occurred only when the gaze shift involved head turns. Eye-gaze shifts of identical amplitude in the absence of a head-movement turned out to be ineffective. However, large-amplitude eye-gaze shifts, when presented at a closer distance simulating near-field social interactions, successfully enhanced luminance change detection, indicating covert shifts of attention. Together, these findings suggest that rhesus monkeys are, in principle, capable of exploiting the spatial information offered by a conspecific's eyes. Yet this ability seems to be primarily engaged under close-range conditions, in line with previous studies (Kano et al., 2022), likely suitable to facilitate the scrutiny of the direction of a conspecific's eyes. Thus, poorer input quality, rather than a fundamental absence of the necessary perceptual capacity, appear to limit rhesus monkeys' ability to respond to isolated eye-gaze cues.

To investigate the influence of gaze cues preceding the luminance change on detection accuracy, we compared performance in congruent and incongruent conditions (i.e., trials in which the demonstrator's gaze was directed toward the target object undergoing a luminance change versus trials in which the gaze was directed toward the distractor). We hypothesized that higher detection accuracy in congruent trials, relative to incongruent trials, would reflect a covert shift of attention in the observer monkey, guided by the gaze cues presented by the avatar (Fig. S2). In incongruent trials, the luminance change was expected to be more frequently missed due to attentional misallocation to the opposite target. A strict central fixation requirement during gaze cue presentation precluded overt attentional shifts via eye movements (Fig. S2). Instead of indicating the demonstrator's gaze direction, the observers had learned to use saccades to indicate the object they judged to have undergone a luminance change. As our experimental design eliminated any reward-based incentive for gaze-following, any performance enhancements in congruent trials would reflect spontaneous covert gaze-following tendencies, persisting despite the absence of obvious benefits for the observer.

The observed gaze-mediated attentional modulation occurred as early as 50 ms relative to the gaze cue, no matter if the gaze shift was based on a head turn or an eye movement from the demonstrator brought into closer view with larger amplitude. Such rapid attentional shifts are consistent with a reflexive mechanism that relies on a few conspicuous exogenous features that may be captured without the need for extensive and therefore time-consuming processing (Corbetta & Shulman, 2002; Duncan & Humphreys, 1989; Posner & Cohen, 1984). This early effect is supported by substantial evidence from studies demonstrating that neural correlates of covert attentional shifts can be detected in the thalamus and early visual cortical areas as early as 50

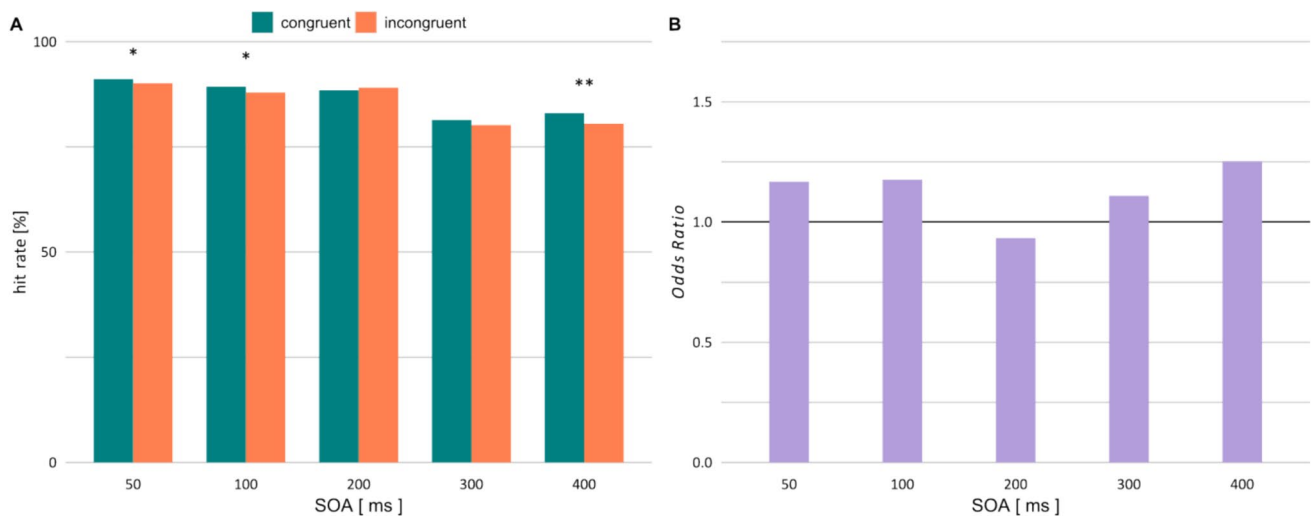


Fig. 4 Hit rate and ORs as a function of SOA for the exaggerated eye-only gaze cue in Monkey E. * $p < .05$; ** $p < .01$; *** $p < .001$. (Color figure online)

ms following cue presentation in nonhuman primates using Posner-type paradigms (Corbetta & Shulman, 2002; McAdams & Reid, 2005; McAlonan et al., 2008). This effect also aligns with the behavioral findings from a previous report showing that rhesus monkeys covertly shift their attention to targets indicated by conspecifics' head-gaze after similarly short latencies (Marciniak et al., 2015). In this study, observers were asked to detect a luminance change of a randomly selected spatial target, to be identified either employing spatial information offered by the other's head gaze or, alternatively, the other's identity based on a learned association between identity and target position. In case the observer was instructed to follow the demonstrator's gaze, a deployment of attention toward the gaze-congruent target was found to start as early as 50 ms and persisted for up to 500 ms. When the prevailing rule required the monkey to ignore the demonstrator's gaze and instead rely on the demonstrator's identity to locate the target, appropriate attentional shifts toward the identity-congruent (and thus correct) target were observed only at longer cue–target intervals (SOAs), emerging around 300 ms. This pattern indicates that the identity-matching rule was successfully learned and applied by the animals. However, under shorter SOAs, beginning around 50 ms, adherence to the rule became more difficult, resulting in frequent errors. Notably, in these error trials, the monkeys consistently selected the gaze-congruent target, suggesting that both the gaze direction and the demonstrator's identity were processed even when the task rule required gaze direction to be ignored. The critical factor determining whether the gaze could be effectively suppressed in order to prioritize the identity cue was therefore the duration for which the demonstrator's image remained visible. This pattern aligns with the notion of limited executive control over an initially

automatic orienting response prompted by a gaze cue. Collectively, these findings suggest that gaze-driven shifts of covert attention comprise two components: an early, largely automatic process that occurs too rapidly to be influenced by top-down control, and a later, more deliberate component that can be modulated by cognitive control mechanisms once sufficient processing time is available (Driver et al., 1999; Friesen & Kingstone, 1998; Frischen et al., 2007; Langton et al., 2000; Marciniak et al., 2015; Ricciardelli et al., 2002; Yoxon et al., 2019).

The early deployment of covert attentional shifts observed in the present study contrasts with the findings of Deaner and Platt (2003), who reported that reflexive, gaze-driven covert shifts of attention did not occur earlier than 200 ms following the presentation of a gaze cue. Aside from the current work, their study remains the only one to directly investigate the influence of conspecifics' eyes, either with or without accompanying head turns demonstrating nonpredictive gaze cues on covert attentional shifts in rhesus monkeys. In their paradigm, static, grey-scale images of a monkey face with averted eye or head-gaze were shown for varying durations (100, 200, 400, and 800 ms), following the disappearance of a central fixation square. Monkeys were then required to make a saccade to a target appearing at either a gaze-congruent or incongruent location. Although their findings generally support the notion that rhesus monkeys can use conspecific eye or head-gaze direction to covertly shift attention, the late onset of the shift seems not easily reconcilable with the demonstration of an early gaze-dependent attention reflex in Marciniak et al. (2015) and in the present study. A possible explanation may be provided by key methodological differences. Whereas Deaner and Platt presented gaze

cues in isolation, without a preceding reference image, both our study and that of Marciniak used a two-step cueing approach: a forward-looking face with direct gaze preceded the gaze cue. We propose that this well-defined transition between direct and averted gaze may render the change in gaze direction, precisely located in time, more salient and easier to detect, thereby accelerating the recognition of behaviorally relevant visual features. Conversely, when a gaze cue is presented without prior direct gaze serving as a baseline, the extraction of critical spatial-cueing information may be delayed. This interpretation is supported by findings from two previous behavioral studies, one in humans and the other in rhesus macaques, which reported a significantly stronger gaze-cueing effect when the cue was preceded by a centrally presented, forward-looking face, compared with trials without such a referential stimulus (Driver et al., 1999; Yu et al., 2012).

A key finding of the present study was the demonstration of covert visual attentional modulation in response to isolated eye-gaze cues, effectively rejecting the hypothesis that rhesus macaques may lack the perceptual capacity to use spatial information provided by the eyes to guide the allocation of attention. Instead, our results highlight the critical influence of two crucial factors in eliciting eye-gaze following behavior: the distance between the demonstrator and the observer, and the amplitude of the eye-gaze cue. When the eye-gaze cue is subtle and observed from a greater distance, it may go unnoticed. In contrast, when the cue is viewed from close proximity and is sufficiently large in amplitude, it effectively elicits gaze-following behavior in rhesus monkeys. This is consistent with the findings of Kano et al. (2022), who demonstrated that in a three-alternative forced-choice task, chimpanzees' accuracy in discriminating a conspecific's averted-gaze face from two direct-gaze faces declined with increasing viewing distance. Because we simultaneously manipulated both proximity and amplitude, it is difficult to determine which factor had the greater impact. In any case, our findings clearly indicate that the ability to follow eye-gaze is present in rhesus monkeys, a species whose evolutionary lineage diverged from that of modern humans more than 30 million years ago (Pozzi et al., 2014). This suggests that the capability to use eye-gaze direction for extracting behaviorally relevant information dates back at least to the last common ancestor of old world monkeys (which include rhesus monkeys) and hominids (i.e., apes and both modern and pre-modern humans). This behavioral finding is in line with neurophysiological data from single-unit recordings in rhesus monkeys, which have identified neurons in the superior temporal sulcus (STS) that are selectively tuned to the direction of eye-gaze, independent of head orientation (Perrett et al., 1985; Yang & Freiwald, 2021). However, it remains unclear to what extent this ability is employed in everyday social interactions or how

effectively eye-gaze cues function for long-range communication in this species.

The prevailing view holds that humans possess a markedly superior capacity for eye-gaze following relative to NHPs. This claim partly rests on a series of studies reporting that eye-gaze following is completely absent in NHPs (Burkart & Heschl, 2006; Ferrari et al., 2000; Peignot & Anderson, 1999; Santos & Hauser, 1999) or is substantially weaker than head-gaze following (Tomasello et al., 2007). Of the twelve investigations on eye-gaze following in NHPs (Burkart & Heschl, 2006; Deaner & Platt, 2003; Ferrari et al., 2000; Itakura & Tanaka, 1998; Kano et al., 2022; Lorincz et al., 2001; Peignot & Anderson, 1999; Povinelli & Eddy, 1996, 1997; Santos & Hauser, 1999; Tomasello et al., 2007; Yu et al., 2012), only five reported negative results, either an absence or weaker eye-gaze following abilities (Burkart & Heschl, 2006; Ferrari et al., 2000; Peignot & Anderson, 1999; Santos & Hauser, 1999). Notably, all five employed live human demonstrators, which raises concerns regarding the ecological validity of human gaze cues for NHPs and whether failure in this task indicates a broader inability to follow gaze from conspecifics. Additionally, these studies lacked high-resolution eye-tracking, relying instead on human-scored video recordings with frame rates insufficient to reliably capture rapid eye movements such as gaze-driven saccades. Some paradigms required gaze-follow into empty space (Ferrari et al., 2000; Tomasello et al., 2007), known to induce habituation over repeated exposure thus attenuating subsequent responses (Parron & Meguerditchian, 2016; Tomasello et al., 2007), as gaze-following is typically object-directed. Finally, inconsistent results across studies might reflect uncontrolled variables, such as the quantity and quality of NHP exposure to humans or from age-related differences in gaze-following propensity. Indeed, increased exposure to humans is known to enhance NHPs' ability to follow human eye-gaze (Byrnit, 2009; Ferrari et al., 2000, 2008; Parron & Meguerditchian, 2016; Peignot & Anderson, 1999), whereas the influence of age remains unclear. Some reports documented improvements in eye-gaze following with age (Ferrari et al., 2000, 2008), while others reported the opposite (Parron & Meguerditchian, 2016; Tomasello et al., 2007). Collectively, the evidence supporting the notion that eye-gaze following is unique to humans is undermined by substantial methodological limitations, casting doubt on the reliability of the conclusion that NHPs lack this capacity.

Given the lack of sound empirical evidence supporting the claim that eye-gaze following behavior is uniquely human, it is crucial to reconsider the second major argument supporting exclusivity of eye-gaze following in humans: the supposed exceptional visibility of human eyes. Early qualitative studies by Kobayashi and Kohshima (1997, 2001), building on Morris (1985), argued that humans possess

uniquely visible eyes, characterized by a pale sclera, a high eye width-to-height ratio (WHR), and substantial scleral exposure. However, these conclusions have been challenged by recent findings on several grounds. First, pale peri-iridal tissues have been documented in various Old World and New World monkey species, and significant variation exists within species concerning scleral pigmentation and eye shape (Caspar et al., 2021; Clark et al., 2023; Mayhew & Gómez, 2015; Perea García, 2016; Perea-García et al., 2019, 2022, 2024). For instance, 17% of Ngogo chimpanzees, despite typically having dark sclera, exhibit pale sclera (Clark et al., 2023), while many humans living in equatorial or rural regions commonly have conjunctival pigmentation (Blake et al., 2003; Jakobieć, 1984, 2016; Mann, 1966; Perea-García et al., 2025; Singh et al., 1998; Whittington et al., 2024). Additionally, WHR varies across humans, and NHPs like orangutans have WHR values comparable to humans (Kaplan & Rogers, 2002). Second, gaze visibility is determined by the contrast between iris, sclera, and peri-orbital skin rather than scleral brightness alone, with studies revealing no significant luminance contrast differences among humans, chimpanzees, and bonobos (Mearing & Koops, 2021; Perea-García et al., 2019; Whitham et al., 2022a, 2022b). Third, early investigations focused on direct gaze when comparing the area of exposed sclera across species, while subsequent studies have shown that, when gaze is averted, the proportion of visible sclera does not differ significantly between humans and gorillas (Mayhew & Gómez, 2015). The contradictory evidence undermines the hypothesis that humans possess uniquely conspicuous ocular morphology. Consequently, it challenges the claim that this trait underpins superior eye-gaze following capabilities, and exposes the limitations of the initial, overly simplistic visibility metrics. Hence, a more nuanced view should be adopted because gaze-cue visibility likely depends on many factors some of which may be interacting such as observer-demonstrator distance, ambient illumination, and the combined hue-luminance contrasts of the iris and sclera. Moreover, the specific contribution of hue contrast among ocular structures remains unexplored; configurations matched for luminance may still differ in hue contrast and, therefore, in perceptual salience (Some of these understudied morphological features, which may contribute to communicative or protective ocular functions, are described in detail by Perea-García et al., 2021).

In summary current evidence, including our own findings, rejects a qualitative gap in eye-gaze following ability between humans and nonhuman primates (NHPs). However, nevertheless there could still be significant quantitative differences. Three comparative studies have directly addressed this question. Two reported markedly larger effects in humans: Tomasello et al. (2007) found that 12- to 18-month-old infants showed 1.26-fold higher eye-gaze

following accuracy than great apes (Cohen's $d = 1.53$ vs. 1.22, respectively), and Deaner and Platt (2003) observed reaction-time and eye-position effects in humans that were 1.7–3.6 times those of monkeys. Conversely, Itakura and Tanaka (1998) detected no species difference in an object-choice task involving chimpanzees, an orangutan, and 23-month-old infants. This limited, methodologically diverse literature tentatively favors greater human proficiency but precludes firm conclusions, highlighting the need for more well-controlled experiments. Such comparative studies should present gaze cues from both human and nonhuman conspecifics enabling a more precise dissection of intra- and inter-specific similarities in gaze-following performance. High-fidelity virtual avatars, like those we incorporated in the present study, provide an optimal platform for this line of research. They enable systematic manipulation of factors influencing gaze cue salience, including demonstrator–observer distance, illumination, scleral and peri-orbital luminance and hue contrasts, eye shape, exposed scleral area, and gaze-motion kinematics, while maintaining ecologically valid stimuli.

In conclusion, the present study demonstrates that the visuomotor system of rhesus monkeys is capable of following eye-gaze, provided that the gaze cue is sufficiently salient, such as when viewed at close proximity with large enough amplitude. The gaze-driven modulation of covert visual spatial attention occurs reflexively in response to salient visual input. Notably, we also observed attentional effects persisting at longer cue-target intervals, indicating the possible engagement of endogenous, goal-directed attention processes in gaze-following. Furthermore, the use of a realistic monkey avatar proved essential for addressing our research question with greater precision which opens promising avenues for future studies exploring the sensory and cognitive mechanisms underlying social cue processing in nonhuman primates and humans.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13415-025-01383-0>.

Acknowledgements We thank Dr. Friedemann Bunjes and Peter W. Dicke for their invaluable technical assistance.

Author contributions M.G. and P.T. contributed to conceptualization; M.R. and M.G. contributed to data curation; M.S. contributed to formal analysis; M.Gi. and P.T. contributed to funding acquisition; M.R. contributed to investigation; M.S., M.G., and P.T. contributed to methodology; M.S. contributed to project administration; M.Gi. and P.T. contributed to resources; M.G. and N.T. contributed to software; P.T. contributed to supervision; M.S. and M.G. contributed to validation; M.S. and M.R. contributed to visualization; M.S. and M.R. contributed to writing—original draft; M.S. and P.T. contributed to writing—review & editing.

Funding Open Access funding enabled and organized by Projekt DEAL. We acknowledge support from the German Research

Foundation (DFG) project TH425/17–1 and from the Open Access Publication Fund of the University of Tübingen.

Data availability The data have been deposited on the Open Science Framework (OSF) and is available upon request. For review purposes, access has been granted through the following exclusive link: https://osf.io/mx2g7/?view_only=00e50c5d3190458ea2d9e1fe3ba3b15e

Code availability Those scripts used for the statistical analysis of the data are publicly available at: https://github.com/shafiei-masih/SocialAttention_monkeys

Declarations

Competing interests The authors declare that they have no conflicts of interest related to this study.

Ethics approval All experimental procedures complied with German and European Guidelines for the Care and Use of Laboratory Animals and were reviewed and approved by the responsible local authorities (Regierungspräsidium Tübingen N07/18 and Landratsamt Tübingen).

Consent to participate Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anderson, J. R., & Mitchell, R. W. (1999). Macaques but not lemurs co-orient visually with humans. *Folia Primatologica*, 70(1), 17–22. <https://doi.org/10.1159/000021670>
- Anderson, N. C., Risko, E. F., & Kingstone, A. (2016). Motion influences gaze direction discrimination and disambiguates contradictory luminance cues. *Psychonomic Bulletin & Review*, 23, 817–823. <https://doi.org/10.3758/s13423-015-0971-8>
- Atabaki, A., Marciniak, K., Dicke, P. W., & Thier, P. (2015). Assessing the precision of gaze following using a stereoscopic 3D virtual reality setting. *Vision Research*, 112, 68–82. <https://doi.org/10.1016/j.visres.2015.04.015>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bechert, K., & Koenig, E. (1996). A search coil system with automatic field stabilization, calibration, and geometric processing for eye movement recording in humans. *Neuro-Ophthalmology*. <https://doi.org/10.3109/01658109609009677>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Beuker, K. T., Rommelse, N. N. J., Donders, R., & Buitelaar, J. K. (2013). Development of early communication skills in the first two years of life. *Infant Behavior & Development*, 36(1), 71–83. <https://doi.org/10.1016/j.infbeh.2012.11.001>
- Blake, C. R., Lai, W. W., & Edward, D. P. (2003). Racial and ethnic differences in ocular anatomy. *International Ophthalmology Clinics*, 43(4), 9.
- Bock, S. W., Dicke, P., & Thier, P. (2008). How precise is gaze following in humans? *Vision Research*, 48(7), 946–957. <https://doi.org/10.1016/j.visres.2008.01.011>
- Bräuer, J., Call, J., & Tomasello, M. (2005). All great ape species follow gaze to distant locations and around barriers. *Journal of Comparative Psychology*, 119(2), 145–154. <https://doi.org/10.1037/0735-7036.119.2.145>
- Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language*, 35(1), 207–220. <https://doi.org/10.1017/S030500090700829X>
- Burkart, J., & Heschl, A. (2006). Geometrical gaze following in common marmosets (*Callithrix jacchus*). *Journal of Comparative Psychology*, 120(2), 120–130. <https://doi.org/10.1037/0735-7036.120.2.120>
- Burnham, K. P., & Anderson, D. R. (Eds.). (2004). *Model selection and multimodel inference*. Springer. <https://doi.org/10.1007/b97636>
- Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9(1), 55–72. <https://doi.org/10.1111/j.2044-835X.1991.tb00862.x>
- Byrmit, J. T. (2009). Gorillas' (*Gorilla gorilla*) use of experimenter-given manual and facial cues in an object-choice task. *Animal Cognition*, 12(2), 401–404. <https://doi.org/10.1007/s10071-008-0200-1>
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51(13), 1484–1525. <https://doi.org/10.1016/j.visres.2011.04.012>
- Caspar, K. R., Biggemann, M., Geissmann, T., & Begall, S. (2021). Ocular pigmentation in humans, great apes, and gibbons is not suggestive of communicative functions. *Scientific Reports*, 11(1), 12994. <https://doi.org/10.1038/s41598-021-92348-z>
- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development*, 15(4), 481–498. [https://doi.org/10.1016/S0885-2014\(01\)00037-5](https://doi.org/10.1016/S0885-2014(01)00037-5)
- Cheney, D., Seyfarth, R., & Smuts, B. (1986). Social relationships and social cognition in nonhuman primates. *Science*, 234(4782), 1361–1366. <https://doi.org/10.1126/science.3538419>
- Clark, I. R., Lee, K. C., Poux, T., Langergraber, K. E., Mitani, J. C., Watts, D., & Sandel, A. A. (2023). White sclera is present in chimpanzees and other mammals. *Journal of Human Evolution*, 176, 103322. <https://doi.org/10.1016/j.jhevol.2022.103322>
- Congiu, S., Fadda, R., Doneddu, G., & Striano, T. (2016). Impaired representational gaze following in children with autism spectrum disorder. *Research in Developmental Disabilities*, 57, 11–17. <https://doi.org/10.1016/j.ridd.2016.06.008>
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3), 201–215. <https://doi.org/10.1038/nrn755>
- Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology*, 40(2), 271–283. <https://doi.org/10.1037/0012-1649.40.2.271>

- Deaner, R. O., & Platt, M. L. (2003). Reflexive social attention in monkeys and humans. *Current Biology*, 13(18), 1609–1613. <https://doi.org/10.1016/j.cub.2003.08.025>
- Drijvers, L., & Holler, J. (2023). The multimodal facilitation effect in human communication. *Psychonomic Bulletin & Review*, 30(2), 792–801. <https://doi.org/10.3758/s13423-022-02178-x>
- Driver, J., IV., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, 6(5), 509–540. <https://doi.org/10.1080/135062899394920>
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96(3), 433–458. <https://doi.org/10.1037/0033-295X.96.3.433>
- Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, T., BASIS Team. (2012). Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Current Biology*, 22(4), 338–342. <https://doi.org/10.1016/j.cub.2011.12.056>
- 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(3), 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(3/4), 239–249. <https://doi.org/10.1080/17470910701429065>
- 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 of the United States of America*, 97(25), 13997–14002. <https://doi.org/10.1073/pnas.250241197>
- Fox, J. (2015). *Applied regression analysis and generalized linear models*. SAGE Publications.
- Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association*, 87(417), 178–183. <https://doi.org/10.1080/01621459.1992.10475190>
- Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). SAGE Publications.
- Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 5(3), 490–495. <https://doi.org/10.3758/BF03208827>
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention. *Psychological Bulletin*, 133(4), 694–724. <https://doi.org/10.1037/0033-2909.133.4.694>
- Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2019). *Cognitive neuroscience: The biology of the mind* (5th ed). W. W. Norton & Company.
- Goossens, B. M. A., Dekleva, M., Reader, S. M., Sterck, E. H. M., & Bolhuis, J. J. (2008). Gaze following in monkeys is modulated by observed facial expressions. *Animal Behaviour*, 75(5), 1673–1681. <https://doi.org/10.1016/j.anbehav.2007.10.020>
- Itakura, S. (1996). An exploratory study of gaze-monitoring in nonhuman primates. *Japanese Psychological Research*, 38(3), 174–180. <https://doi.org/10.1111/j.1468-5884.1996.tb00022.x>
- Itakura, S., & Tanaka, M. (1998). Use of experimenter-given cues during object-choice tasks by chimpanzees (*Pan troglodytes*), an orangutan (*Pongo pygmaeus*), and human infants (*Homo sapiens*). *Journal of Comparative Psychology*, 112(2), 119–126. <https://doi.org/10.1037/0735-7036.112.2.119>
- Jakobiec, F. A. (1984). The ultrastructure of conjunctival melanocytic tumors. *Transactions of the American Ophthalmological Society*, 82, 599–752.
- Jakobiec, F. A. (2016). Conjunctival primary acquired melanosis: Is it time for a new terminology? *American Journal of Ophthalmology*, 162, 3–19.e1. <https://doi.org/10.1016/j.ajo.2015.11.003>
- Kano, F., & Call, J. (2014). Cross-species variation in gaze following and conspecific preference among great apes, human infants and adults. *Animal Behaviour*, 91, 137–150. <https://doi.org/10.1016/j.anbehav.2014.03.011>
- Kano, F., Kawaguchi, Y., & Hanling, Y. (2022). Experimental evidence that uniformly white sclera enhances the visibility of eye-gaze direction in humans and chimpanzees. *eLife*, 11, e74086. <https://doi.org/10.7554/eLife.74086>
- Kaplan, G., & Rogers, L. J. (2002). Patterns of gazing in orangutans (*Pongo pygmaeus*). *International Journal of Primatology*, 23(3), 501–526. <https://doi.org/10.1023/A:1014913532057>
- Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature*, 387(6635), 767–768. <https://doi.org/10.1038/42842>
- Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: Comparative studies on external morphology of the primate eye. *Journal of Human Evolution*, 40(5), 419–435. <https://doi.org/10.1006/JHEV.2001.0468>
- Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, 4(2), 50–59. [https://doi.org/10.1016/S1364-6613\(99\)01436-9](https://doi.org/10.1016/S1364-6613(99)01436-9)
- Lenth, R. V. (2025). *emmeans: Estimated marginal means, aka least-squares means* (Version 1.10.7) [Computer software]. <https://CRAN.R-project.org/package=emmeans>
- Liebal, K., & Kaminski, J. (2012). Gibbons (*Hylobates pileatus*, *H. moloch*, *H. lar*, *Symphalangus syndactylus*) follow human gaze, but do not take the visual perspective of others. *Animal Cognition*, 15(6), 1211–1216. <https://doi.org/10.1007/s10071-012-0543-5>
- Lorincz, E. N., Baker, C. I., & Perrett, D. I. (2001). Visual cues for attention following in rhesus monkeys. In J. Fagot (Ed.), *Picture perception in animals* (pp. 343–372). Psychology Press.
- Lu, X., Wang, Q., Li, X., Wang, G., Chen, Y., Li, X., & Li, H. (2023). Connectivity reveals homology between the visual systems of the human and macaque brains. *Frontiers in Neuroscience*. <https://doi.org/10.3389/fnins.2023.1207340>
- Lüdtke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *The Journal of Open Source Software*, 6(60), 3139. <https://doi.org/10.21105/joss.03139>
- Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, 49(4), 1494–1502. <https://doi.org/10.3758/s13428-016-0809-y>
- Mann, I. (1966). *Culture, race, climate and eye disease: An introduction to the study of geographic ophthalmology*. Thomas.
- Marciniak, K., Dicke, P. W., & Thier, P. (2015). Monkeys head-gaze following is fast, precise and not fully suppressible. *Proceedings of the Royal Society B: Biological Sciences*, 282(1816), 20151020. <https://doi.org/10.1098/rspb.2015.1020>
- Mayhew, J. A., & Gómez, J.-C. (2015). Gorillas with white sclera: A naturally occurring variation in a morphological trait linked to social cognitive functions. *American Journal of Primatology*, 77(8), 869–877. <https://doi.org/10.1002/ajp.22411>
- McAdams, C. J., & Reid, R. C. (2005). Attention modulates the responses of simple cells in monkey primary visual cortex. *Journal of Neuroscience*, 25(47), 11023–11033. <https://doi.org/10.1523/JNEUROSCI.2904-05.2005>
- McAlonan, K., Cavanaugh, J., & Wurtz, R. H. (2008). Guarding the gateway to cortex with attention in visual thalamus. *Nature*, 456(7220), 391–394. <https://doi.org/10.1038/nature07382>
- Mearing, A. S., & Koops, K. (2021). Quantifying gaze conspicuity: Are humans distinct from chimpanzees and bonobos? *Journal of Human Evolution*, 157, 103043. <https://doi.org/10.1016/j.jhevol.2021.103043>

- Mehrabian, A. (2017). Nonverbal communication. *Routledge*. <https://doi.org/10.4324/9781351308724>
- Moors, P., Germeyns, F., Pomianowska, I., & Verfaillie, K. (2015). Perceiving where another person is looking: The integration of head and body information in estimating another person's gaze. *Frontiers in Psychology*, 6, 909. <https://doi.org/10.3389/fpsyg.2015.00909>
- Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development*, 21(2), 373–377. [https://doi.org/10.1016/S0163-6383\(98\)90014-5](https://doi.org/10.1016/S0163-6383(98)90014-5)
- Morioka, S., Osumi, M., Shiotani, M., Nobusako, S., Maeoka, H., Okada, Y., Hiyamizu, M., & Matsuo, A. (2016). Incongruence between verbal and nonverbal information enhances the late positive potential. *PLoS ONE*, 11(10), Article e0164633. <https://doi.org/10.1371/journal.pone.0164633>
- Morris, D. (1985). *Bodywatching: A field guide to the human species*. Crown.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>
- Parron, C., & Meguerditchian, A. (2016). Gaze following in baboons (*Papio anubis*): Juveniles adjust their gaze and body position to human's head redirections. *American Journal of Primatology*, 78(12), 1265–1271. <https://doi.org/10.1002/ajp.22580>
- Peignot, P., & Anderson, J. R. (1999). Use of experimenter-given manual and facial cues by gorillas (*Gorilla gorilla*) in an object-choice task. *Journal of Comparative Psychology*, 113(3), 253–260. <https://doi.org/10.1037/0735-7036.113.3.253>
- Perea García, J. O. (2016). Quantifying ocular morphologies in extant primates for reliable interspecific comparisons. *Journal of Language Evolution*, 1(2), 151–158. <https://doi.org/10.1093/jole/lzw004>
- Perea-García, J. O., Danel, D. P., & Monteiro, A. (2021). Diversity in primate external eye morphology: Previously undescribed traits and their potential adaptive value. *Symmetry*, 13(7), 1270.
- Perea-García, J. O., Kret, M. E., Monteiro, A., & Hobaiter, C. (2019). Scleral pigmentation leads to conspicuous, not cryptic, eye morphology in chimpanzees. *Proceedings of the National Academy of Sciences of the United States of America*, 116(39), 19248–19250. <https://doi.org/10.1073/pnas.1911410116>
- Perea-García, J. O., Massen, J. J. M., Ostner, J., Schülke, O., Castellano-Navarro, A., Gazagne, E., & Monteiro, A. (2024). Photoregulatory functions drive variation in eye coloration across macaque species. *Scientific Reports*, 14(1), 29115. <https://doi.org/10.1038/s41598-024-80643-4>
- Perea-García, J. O., Ramarajan, K., Kret, M. E., Hobaiter, C., & Monteiro, A. (2022). Ecological factors are likely drivers of eye shape and colour pattern variations across anthropoid primates. *Scientific Reports*, 12(1), 17240. <https://doi.org/10.1038/s41598-022-20900-6>
- Perea-García, J. O., Teuben, A., & Caspar, K. R. (2025). Look past the cooperative eye hypothesis: Reconsidering the evolution of human eye appearance. *Biological Reviews*. <https://doi.org/10.1111/brv.70033>
- Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., . . . Weiskrantz, L. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 223(1232), 293–317. <https://doi.org/10.1098/rspb.1985.0003>
- Platt, M. L., Seyfarth, R. M., & Cheney, D. L. (2016). Adaptations for social cognition in the primate brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1687), 20150096. <https://doi.org/10.1098/rstb.2015.0096>
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25. <https://doi.org/10.1080/00335558008248231>
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531–556). Erlbaum.
- Povinelli, D. J., & Eddy, T. J. (1996). Chimpanzees: Joint visual attention. *Psychological Science*, 7(3), 129–135. <https://doi.org/10.1111/j.1467-9280.1996.tb00345.x>
- Povinelli, D. J., & Eddy, T. J. (1997). Specificity of gaze-following in young chimpanzees. *British Journal of Developmental Psychology*, 15(2), 213–222. <https://doi.org/10.1111/j.2044-835X.1997.tb00735.x>
- Pozzi, L., Hodgson, J. A., Burrell, A. S., Sterner, K. N., Raaum, R. L., & Disotell, T. R. (2014). Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. *Molecular Phylogenetics and Evolution*, 75, 165–183. <https://doi.org/10.1016/j.ympev.2014.02.023>
- Prsa, M., & Thier, P. (2011). The role of the cerebellum in saccadic adaptation as a window into neural mechanisms of motor learning. *European Journal of Neuroscience*, 33(11), 2114–2128. <https://doi.org/10.1111/j.1460-9568.2011.07693.x>
- R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ricciardelli, P., Bricolo, E., Aglioti, S. M., & Chelazzi, L. (2002). My eyes want to look where your eyes are looking: Exploring the tendency to imitate another individual's gaze. *NeuroReport*, 13(17), 2259.
- 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: Biological Sciences*, 283(1830), 20160376. <https://doi.org/10.1098/rspb.2016.0376>
- Ruiz, A., Gómez, J. C., Roeder, J. J., & Byrne, R. W. (2009). Gaze following and gaze priming in lemurs. *Animal Cognition*, 12(3), 427–434. <https://doi.org/10.1007/s10071-008-0202-z>
- Santos, L. R., & Hauser, M. D. (1999). How monkeys see the eyes: Cotton-top tamarins' reaction to changes in visual attention and action. *Animal Cognition*, 2(3), 131–139. <https://doi.org/10.1007/s100710050033>
- Shepherd, S. V., & Platt, M. L. (2008). Spontaneous social orienting and gaze following in ringtailed lemurs (*Lemur catta*). *Animal Cognition*, 11(1), 13–20. <https://doi.org/10.1007/s10071-007-0083-6>
- Siebert, R., Taubert, N., Spadacenta, S., Dicke, P. W., Giese, M. A., & Thier, P. (2020). A naturalistic dynamic monkey head avatar elicits species-typical reactions and overcomes the uncanny valley. *ENeuro*. <https://doi.org/10.1523/ENEURO.0524-19.2020>
- Singh, A. D., Campos, O. E., Rhatigan, R. M., Schulman, J. A., & Misra, R. P. (1998). Conjunctival melanoma in the Black population. *Survey of Ophthalmology*, 43(2), 127–133. [https://doi.org/10.1016/S0039-6257\(98\)00020-4](https://doi.org/10.1016/S0039-6257(98)00020-4)
- Sodian, B., & Kristen-Antonow, S. (2015). Declarative joint attention as a foundation of theory of mind. *Developmental Psychology*, 51(9), 1190–1200. <https://doi.org/10.1037/dev0000039>
- Symonds, M. R. E., & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology*, 65(1), 13–21. <https://doi.org/10.1007/s00265-010-1037-6>
- Teufel, C., Gutmann, A., Pirow, R., & Fischer, J. (2010). Facial expressions modulate the ontogenetic trajectory of gaze-following among monkeys. *Developmental Science*, 13(6), 913–922. <https://doi.org/10.1111/j.1467-7687.2010.00956.x>

- Thorup, E., Nyström, P., Gredebäck, G., Bölte, S., Falck-Ytter, T., EASE Team. (2016). Altered gaze following during live interaction in infants at risk for autism: An eye tracking study. *Molecular Autism*, 7, Article 12. <https://doi.org/10.1186/s13229-016-0069-9>
- Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, 55(4), 1063–1069. <https://doi.org/10.1006/anbe.1997.0636>
- Tomasello, M., Hare, B., & Agnetta, B. (1999). Chimpanzees, *Pan troglodytes*, follow gaze direction geometrically. *Animal Behaviour*, 58(4), 769–777. <https://doi.org/10.1006/anbe.1999.1192>
- Tomasello, M., Hare, B., & Fogleman, T. (2001). The ontogeny of gaze following in chimpanzees, *Pan troglodytes*, and rhesus macaques, *Macaca mulatta*. *Animal Behaviour*, 61(2), 335–343. <https://doi.org/10.1006/anbe.2000.1598>
- Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution*, 52(3), 314–320. <https://doi.org/10.1016/j.jhevol.2006.10.001>
- Whitham, W., Schapiro, S. J., Troscianko, J., & Yorzinski, J. L. (2022a). Chimpanzee (*Pan troglodytes*) gaze is conspicuous at ecologically-relevant distances. *Scientific Reports*, 12(1), 9249. <https://doi.org/10.1038/s41598-022-13273-3>
- Whitham, W., Schapiro, S. J., Troscianko, J., & Yorzinski, J. L. (2022b). The gaze of a social monkey is perceptible to conspecifics and predators but not prey. *Proceedings of the Royal Society b: Biological Sciences*, 289(1976), 20220194. <https://doi.org/10.1098/rspb.2022.0194>
- Whittington, C. P., Bresler, S. C., Simon, C., Shields, C. L., & Patel, R. M. (2024). Melanocytic lesions of the conjunctiva: An up-to-date review. *Diagnostic Histopathology*, 30(1), 37–59. <https://doi.org/10.1016/j.mpdhp.2023.10.005>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.
- Yang, Z., & Freiwald, W. A. (2021). Joint encoding of facial identity, orientation, gaze, and expression in the middle dorsal face area. *Proceedings of the National Academy of Sciences of the United States of America*, 118(33), Article e2108283118. <https://doi.org/10.1073/pnas.2108283118>
- Yoxon, E., Constable, M. D., & Welsh, T. N. (2019). Probing the time course of facilitation and inhibition in gaze cueing of attention in an upper-limb reaching task. *Attention, Perception & Psychophysics*, 81(7), 2410–2423. <https://doi.org/10.3758/s13414-019-01821-5>
- Yu, D., Teichert, T., & Ferrera, V. P. (2012). Orienting of attention to gaze direction cues in rhesus macaques: Species-specificity, and effects of cue motion and reward predictiveness. *Frontiers in Psychology*, 3, 202. <https://doi.org/10.3389/fpsyg.2012.00202>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.