



# Orienting of attention to gaze direction cues in rhesus macaques: species-specificity, and effects of cue motion and reward predictiveness

Dian Yu<sup>1</sup>, Tobias Teichert<sup>1</sup> and Vincent P. Ferrera<sup>1,2\*</sup>

<sup>1</sup> Department of Neuroscience, Mahoney Center, Columbia University, New York, NY, USA

<sup>2</sup> Department of Psychiatry, Mahoney Center, Columbia University, New York, NY, USA

## Edited by:

David A. Washburn, Georgia State University, USA

## Reviewed by:

Lisa A. Parr, Emory University, USA  
Francine L. Dolins, University of Michigan-Dearborn, USA  
Federico Sanabria, Arizona State University, USA

## \*Correspondence:

Vincent P. Ferrera, Departments of Neuroscience and Psychiatry, Columbia University, 1051 Riverside Drive, New York, NY 10032, USA.  
e-mail: vpf3@columbia.edu

Primates live in complex social groups and rely on social cues to direct their attention. For example, primates react faster to an unpredictable stimulus after seeing a conspecific looking in the direction of that stimulus. In the current study we tested the specificity of facial cues (gaze direction) for orienting attention and their interaction with other cues that are known to guide attention. In particular, we tested whether macaque monkeys only respond to gaze cues from conspecifics or if the effect generalizes across species. We found an attentional advantage of conspecific faces over human and cartoon faces. Because gaze cues are often conveyed by gesture, we also explored the effect of image motion (a simulated glance) on the orienting of attention in monkeys. We found that the simulated glance did not significantly enhance the speed of orienting for monkey-face stimuli, but had a significant effect for images of human faces. Finally, because gaze cues presumably guide attention toward relevant or rewarding stimuli, we explored whether orienting of attention was modulated by reward predictiveness. When the cue predicted reward location, face, and non-face cues were effective in speeding responses toward the cued location. This effect was strongest for conspecific faces. In sum, our results suggest that while conspecific gaze cues activate an intrinsic process that reflexively directs spatial attention, its effect is relatively small in comparison to other features including motion and reward predictiveness. It is possible that gaze cues are more important for decision-making and voluntary orienting than for reflexive orienting.

**Keywords:** monkey, conspecific, reflexive, learned, Posner, endogenous, decision-making

## INTRODUCTION

The social orienting hypothesis is the idea that social animals take cues about where to direct their attention from other members of their social group. It posits a natural link between social cues and visual spatial attention. Behavioral correlates of social orienting have been reported in humans and non-human primates (Deaner and Platt, 2003; Nummenmaa and Calder, 2009; Hill et al., 2010). Faces are a natural source of attentional cues that seem to be inherently social. However, gaze direction is a complex visual cue and the specifically social component may be difficult to isolate. Given that animals are more likely to engage in social interactions within their own species, one might predict that if gaze is inherently social, then attention to gaze cues in monkeys should depend on whether the cue is conveyed by another monkey, a human, or other face-like stimulus. In the current study we tested whether orienting of attention to gaze direction in macaques is restricted to cues of conspecifics or if it generalizes to other species or even cartoon drawings of faces. Attention to facial cues may also depend on how realistic and reliable the cue is. Thus, we also tested whether gaze cues are enhanced by adding motion to the cue stimulus, or by making the cue more reliably predictive of reward.

Spatial attention is automatically guided toward salient and or unpredictable stimuli such as a flash of light or a loud noise (reflexive or exogenous orienting of attention). However, less salient cues such as arrows may also orient attention if they reliably predict the occurrence of a relevant stimulus at a particular location in space (endogenous orienting of attention; Jonides, 1981; Nakayama and Mackeben, 1989). Orienting to gaze cues may be an automated or reflexive form of endogenous orienting: it assumes that attention is guided by an innate tendency to orient to stimuli that are being attended by conspecifics. As direction of gaze is highly correlated with the locus of attention (Baron-Cohen, 1995), gaze-following is thought to be an important indicator of social orienting.

Orienting of attention to gaze direction has been studied using an endogenous version of the Posner paradigm (Posner, 1980). In this task, subjects view a picture of a face with gaze averted toward one side, giving the impression that the depicted subject is fixating an object of interest located somewhere off to the side. Human subjects respond significantly faster when asked to make a saccade in same compared to the opposite direction that the depicted subject is attending to (Driver et al., 1999; Friesen and Kingstone, 2003; Friesen et al., 2004; Ristic and Kingstone, 2005; Ristic et al.,

2007; Tipples, 2008). It is important to note that this orienting effect can be considered automatic or reflexive because it reliably occurs in situations when direction of gaze has no predictive value (Friesen and Kingstone, 1998; Driver et al., 1999; Langton and Bruce, 1999; Langton, 2000; Deaner and Platt, 2003).

Similar effects of orienting to gaze direction have also been reported in non-human primates. Face processing and gaze-following behavior in monkeys are special perceptual processes that are innate, and infant monkeys reared with no exposure to any faces are able to discriminate human faces as well as monkey faces (Sugita, 2008). Gaze-following in Barbary monkeys develops within the first year of life with a rapid increase between the fifth and sixth month. Such gaze-following behavior was shown to enhance when accompanied by a facial expression (Teufel et al., 2010). The effect of social cues on choice behavior has been studied in chimpanzees (Horner et al., 2011; Martin et al., 2011). Face recognition and responses to facial expression develop within the first 2 months in infant chimpanzees, but learning of gestural cues such as pointing may take much longer in this species (Tomonaga et al., 2004).

In rhesus monkeys (*Macaca mulatta*) Deaner and Platt (2003) have shown that the direction of gaze or orientation of the head elicit reflexive orienting of attention. These studies suggest that there is a similar neural circuitry mediating social attention in humans and macaques. However, several important questions regarding the precise role of gaze cues in guiding attention remain open. For example, it is not known whether macaques use gaze cues only from conspecifics or if the effect generalizes to other species such as humans. Similarly it is not known whether the effects of gaze cues are uniquely social or if similar effects can be elicited with non-face cues that presumably have little or no social meaning for macaques. Also, in real life, gaze cues are never encountered as static faces images, but are conveyed by motion (gesture) such as a glance. Finally, the value of gaze cues is related to their ability to signal useful information such as reward prediction. To date the interactions between gaze cues and motion or prediction of reward have not been investigated.

In our study, we explored these open questions through two psychophysical experiments, the Reflexive Orienting (RO) task and the Learned Orienting (LO) task. In the RO task, we used static as well as dynamic versions of face and non-face stimuli to study the interaction between gaze direction cues and motion. In the LO task, we investigated whether reward predictability differentially affects orienting to both face and non-face cues. In contrast to the RO task, cue stimuli in the LO task reliably indicated the location of the rewarded target. Orienting to predictive cues in a social setting could allow additional flexibility, such as knowing when it is appropriate to look toward a peripheral target or look away (Hill et al., 2010). We were interested to what degree monkeys voluntarily orient to cues from other monkeys in comparison to humans, cartoons, or non-face spatial cues. Consistent with previous work (Deaner and Platt, 2003), static images of conspecific faces elicited a reflexive orienting effect in monkeys. This effect was not present for human or cartoon-face cues, suggesting an intrinsic species-specific attention orienting mechanism. When dynamic images were used in the RO task, a congruence effect was observed across all face and non-face stimuli presented. However, motion cues had

little to no additional effect on the magnitude of the congruence effect for the conspecific faces. In the LO task, results indicated that the learned component of orienting was not any greater for face cues than non-face cues. However, the largest improvements in accuracy and choice reaction time, behavioral outputs associated with decision-making, were obtained with monkey-face cues. We suggest that gaze cues may be equally if not more important for decision-making than for reflexive orienting of attention.

## MATERIALS AND METHODS

Experiments were performed on three adult male rhesus monkeys (*M. mulatta*) weighing between 7 and 12 kg. All methods were approved by the Institutional Animal Care and Use Committee at Columbia University and the New York State Psychiatric Institute, and adhere to the ARVO Statement for the Use of Animals in Ophthalmic and Visual Research. Monkeys were housed in a room with 20–25 other monkeys. They were able to interact and socialize with a partner animal directly or through a transparent plastic separator. They interacted directly with humans on a daily basis. Monkeys were prepared for experiments by surgical implantation of a post used for head restraint. Eye position was recorded using a monocular scleral search coil system. All surgical procedures were performed using aseptic technique and general (isoflurane 1–3%) anesthesia. Monkeys selected for this behavioral study also participated in single-cell recordings. Even though less invasive eye tracking procedures were available, the monkeys were already implanted with eye coil and headpost prior to this study for physiology experiments. Monkeys were trained to sit in a primate chair for the duration of the experiment with their heads restrained and perform the behavioral tasks. Monkeys were water restricted and received liquids for performing the behavior tasks. Monkeys received water or tang according to preference. For every correct trial, subject received 0.2 ml of juice for performing a trial. The total amount of water was not limited by the experimenters. Monkeys were allowed to continue to receive fluids as long as they were willing to perform the task. Each monkey completed between 800 and 1200 trials per session and received about 150–300 ml per session. Monkeys generally received supplemental fluids or fruit after each session.

## VISUAL STIMULATION AND EYE MOVEMENT RECORDINGS

Visual stimuli were generated and controlled by a Cambridge Research Systems VSG2/5 video frame buffer. The output from the video board was displayed on a calibrated 37-in color monitor (Mitsubishi) with a 60-Hz non-interlaced refresh rate. The monitor stood at a viewing distance of 24 in so that the display area subtended roughly 40° horizontally by 30° vertically. The spatial resolution of the display was 1280 pixels by 1024 lines. The visual stimuli used during the task consisted of a 0.5° square red fixation cue, a 15.2° × 12.2° stimulus, followed by one or two 1.0° circular red saccade target 14.3° from the center. All images used in the tasks (monkey images: Yerkes Regional Primate Center, human images: Sanderson and Paliwal, 2002) were edited and processed in Photoshop CS2 (Adobe). All stimuli were presented on a uniform black background. The frame buffer was programmed to send out digital pulses (frame sync) for timing purposes at the beginning of each video frame in which a target was turned on or off. These

pulses were recorded by the computer using a hardware timer and stored together with the neuronal and eye movement data.

Eye position was monitored using a monocular scleral search coil system (CNC Engineering). The eye position signals were then digitally sampled by computer at 500 Hz per channel, digitized with 12-bit resolution and stored on a disk for offline analysis. Eye velocity was computed from eye-position using a differentiating digital filter (Gaussian first derivative). Eye position and velocity were used to estimate saccade latency, amplitude, and velocity. Saccade onsets and offsets were computed using an acceleration criterion.

## BEHAVIORAL PARADIGMS

Three monkeys (monkeys F, C, and H) were tested in a Reflexive Orienting (RO) task (**Figure 1B**) and two monkeys (monkeys F and C) in the Learned Orienting (LO) task (**Figure 1C**) separately. For both tasks, stimulus image categories included face or face-like cues (a monkey, human, or cartoon face looking to either the right or left) and clearly non-facial cues (an asymmetric form such as an oval with an off-center dot; **Figure 1A**). Each stimulus category consisted of 6–10 different images. The RO task employed both static and dynamic cue stimuli, whereas the LO task consisted of only dynamic cues. Static cues were photographs of faces looking to the left or right, or ovals with an off-center dot. Dynamic cues consisted of two-frame motion. In the first frame, the face was looking straight ahead (or the dot was centered in the oval). In the second frame, the eyes looked to the left or right (or the dot in the oval was offset). The change in horizontal position of the eyes or dot was the same ( $1^\circ$ ) across all stimuli.

The RO task tested whether monkeys would orient toward the direction implied by a cue stimulus, even though this cue did not predict the location of the saccade targets. In the static version of this task (**Figure 1A**), each trial was initiated when the monkey fixated on a small red square ( $0.5^\circ$  width) in the center of the video display presented on a dark background. The monkey continued to look at the fixation stimulus for a delay period of 500–2000 ms (delay times were drawn from a uniform random distribution). After the delay, the cue stimulus (face with averted gaze or off-center dot) was presented in the center of the display while the monkey continued to fixate there. After a further 50- to 550-ms delay (random uniform distribution), a single peripheral saccade target appeared either to the left or right of the cue (the target size was  $1^\circ$ , and the horizontal location was  $14.3^\circ$  from the center of the display). The location of the peripheral target was randomized as was the direction of gaze or dot offset of the cue stimulus. Hence, the cue was not predictive of the target location. If the target position happened to be consistent with the cue direction, the trial was classified as congruent; otherwise the trial was classified incongruent.

In the dynamic version of the RO and LO tasks, a cue stimulus was presented during the initial delay period of 500–2000 ms. This stimulus was either a face with eyes directed straight ahead or a centered dot. The red fixation target was superimposed on the cue image. At the end of the first delay, the cue image was replaced with an image that was identical except that the eyes or dot were offset to the left or right. This created the two-frame motion mentioned above. Except for the appearance of the cue

in the first delay interval, the static and dynamic cue trials were identical.

The monkeys were rewarded with juice or water for making a saccade to the peripheral target. If eye position moved outside the fixation window before the peripheral target appeared, the trial was aborted without reward.

For the LO task, the timing of the fixation spot, stimulus cue, and peripheral target was identical to the RO task. In this task (**Figure 1C**), a single peripheral target was presented on 30% of trials and the location of this target was not predicted by the cue stimulus. These trials were identical to the dynamic version of the RO task, and served as control trials in the LO task. However, on the remaining 70% of trials, two identical peripheral stimuli were presented. One stimulus (the “target”) was congruent with the cue, and the other (the “distractor”) was presented on the opposite side. On these trials, the cue was always congruent with the target, and the monkey was rewarded only if he looked at the target. Unlike the RO task, only dynamic stimuli were used. Single (0 distractor) and double (1 distractor) target trials were randomly interleaved in the ratio 3:7. Single and double target trials were indistinguishable to the subject until the targets appeared.

## STIMULUS SALIENCE

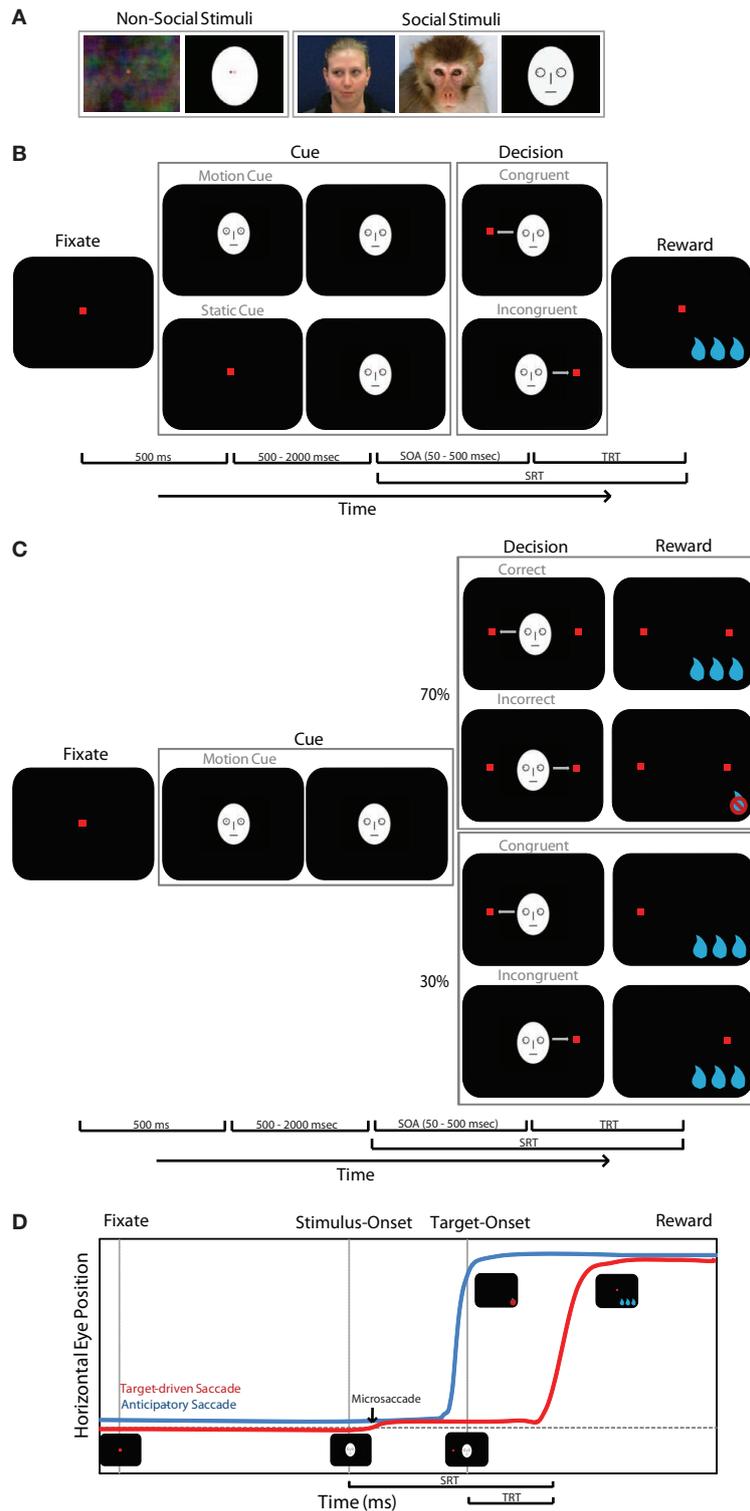
The different classes of image used in these experiments varied not only in terms of high-level content, but also with respect to low level image characteristics, specifically, spatial frequency content and motion energy. To explore whether these image properties might have contributed to the behavioral effects described above, we quantified the spectral energy content of the five classes of image: cartoon face (Cface), static dot on white background (Dots1), moving dot on textured background (Dots2), human face (Hface), and monkey face (Mface). **Figure 2A** shows the Fourier frequency spectrum of each image. For each spatial frequency, energy was averaged over all orientations.

For dynamic stimuli, we computed motion energy (**Figure 2B**) using a direct implementation of the model described by Adelson and Bergen (1985). The parameters of the model were optimized for each image class individually. Of all stimuli used, cartoon-face cues and spatial dot cues possessed the strongest motion energy due to the fact that they had the highest contrast features.

## DATA ANALYSIS

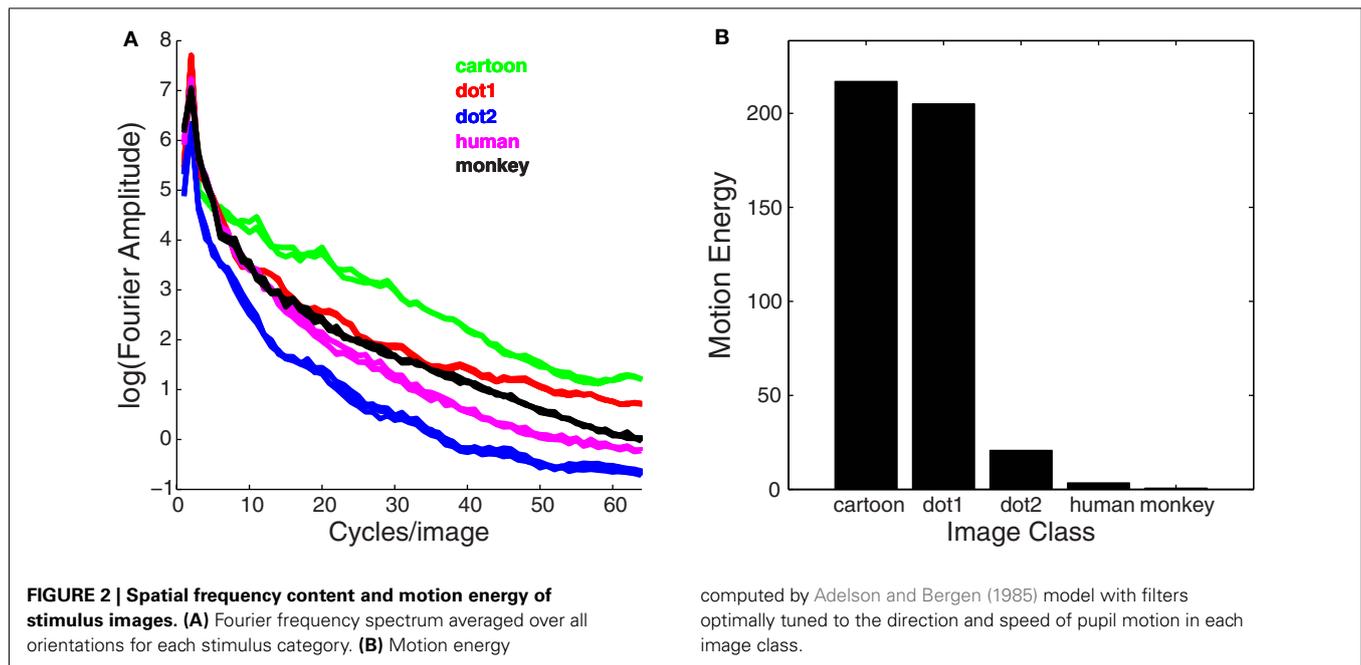
The RO and LO tasks had 32 conditions [4 stimulus types (monkey face, human face, cartoon face, off-center dot)  $\times$  2 cue motion conditions (static or dynamic)  $\times$  2 gaze directions (left/right)  $\times$  2 target directions (left/right)]. Trials were classified as congruent if the gaze direction of the cue image was the same as the target direction, and incongruent if the target direction was opposite to the cue direction. Static and dynamic versions of the RO task were collected in separate alternating sessions. The 0 distractor and 1 distractor trials of the LO task were randomly interleaved within a session.

Saccades in both tasks were quantified as horizontal eye displacements with a minimum amplitude of  $7.15^\circ$  (i.e., saccades that stopped less than half way to the target were excluded). Saccades were divided into three categories: anticipatory saccades, microsaccades, and reactive saccades (**Figure 1D**). Reactive



**FIGURE 1 | Experimental paradigm. (A)** Examples of stimuli used. **(B)** Schematics of the Reflexive Orienting (RO) Task. Animals directed their gaze to the fixation spot, and after 500 ms delay, static or dynamic cue stimuli were shown at random. After a variable SOA, a peripheral target appeared either on the cued or un-cued one side of the stimulus. Animals made saccades to the target to receive reward. **(C)** Schematics of the Learned Orienting (RO) Task. Only dynamic cues

were used for the LO task. Thirty percent of the trials consisted of single target trials, where the center cue had no predictive value regarding target location. The other 70% of trials were dual target trials in which the central cue correctly predicted reward location. Saccades made to the distractor target were not rewarded and counted as incorrect trials. **(D)** RT was calculated as the time from target-onset (not cue onset) to the initiation of the saccade.



saccades were defined as those initiated after the target-onset, and were rewarded if they landed within a  $6^\circ$  window centered on the peripheral target. Anticipatory saccades were those initiated prior to target-onset. Anticipatory saccades accounted for less than 5% of all saccades. Microsaccades were defined as those with horizontal displacements less than  $1.0^\circ$ . Anticipatory saccades and microsaccades were not rewarded.

Reaction time (RT) was measured relative to the onset of the peripheral target, not the onset of the cue. We define cue onset as the time when the face cue appeared for static trials, or the time when the face cue shifted gaze for dynamic trials. The asynchrony between cue and target-onsets varied randomly from 50 to 500 ms. Because shifts of attention are assumed to be triggered by the cue and tend to be transient, it is typical to measure attentional effects as a function of “stimulus onset asynchrony” (here, the difference between cue and target-onset) and then restrict analysis to a range of SOAs that have significant effects. In this study, we combined data across all SOAs.

Prior to statistical tests, RTs were normalized in the following manner: First, trials were grouped by monkey, session, saccade direction, stimulus category, cue type (static or dynamic), number of distractors, and saccade type (anticipatory vs. target-evoked), but not by congruence. The mean RT was calculated within each group. The appropriate mean RT was subtracted from the RT on each trial. Only reactive saccades were used to calculate the congruence effect. Normalized mean RTs were combined across monkeys and sessions, but weighted due to the fact that the two monkeys had unequal amount of sessions and data. A modified version of the weighted two-tailed *t*-test was used to test for significance, where weighted mean and weighted variance replaced mean and variance in calculating *t*-value. ANOVA was used to test the effects of stimulus category on un-normalized RT. A binominal test was used to quantify whether the number of microsaccades and anticipatory saccades made in the direction of the cue was above chance.

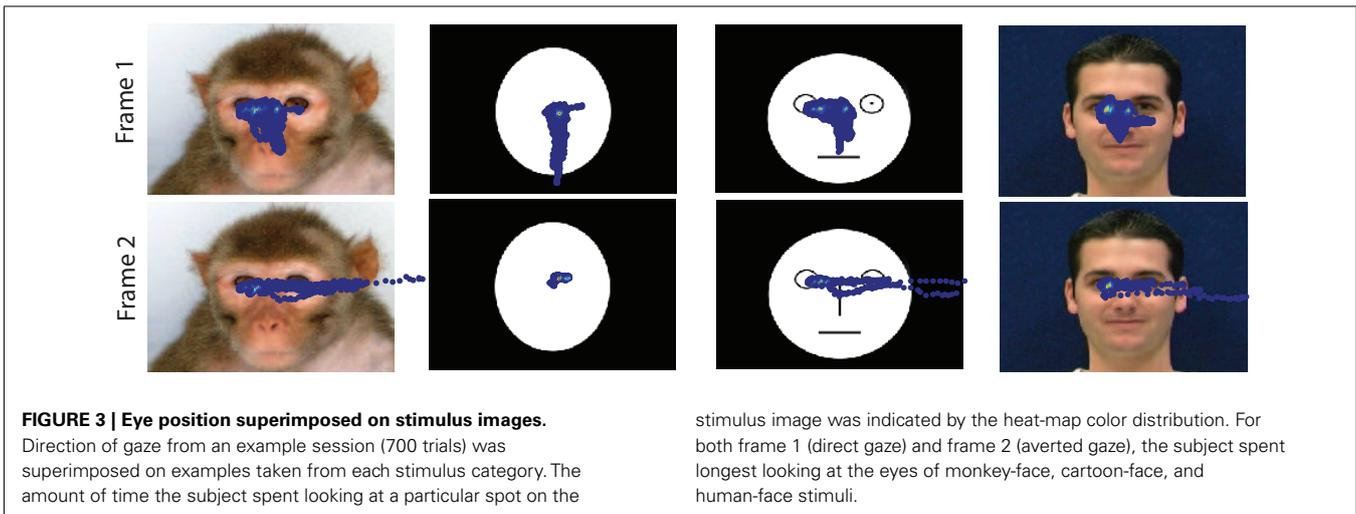
## RESULTS

### REFLEXIVE ORIENTING TASK

Examples of single trial eye movement behavior superimposed on the stimulus images are shown in **Figure 3**. The red fixation target was either centered between the eyes of the face cues or directly overlapping the dot cues. The fixation window was large enough to allow the subjects to explore the entire face image. Subjects preferentially looked at the eyes prior to making saccades directed toward the peripheral target.

We examined the effect of congruence by looking at RT differences between congruent and incongruent trials. The congruence effect has been suggested to reflect intrinsic or reflexive orienting of attention in the direction indicated by the gaze cue (Deaner and Platt, 2003). The number of sessions and trials for each monkey were as follows: Monkey C: 7 sessions, 5363 trials; Monkey F: 13 sessions, 8970 trials; Monkey H: 10 sessions, 9140 trials. Data were normalized by grouping trials by monkey, session, saccade direction, and stimulus type. The mean RT within each group was subtracted from sample RT on each trial. The results are shown in **Figure 4A**. A significant main effect for congruence (congruent vs. incongruent trials) was observed (ANOVA, main effect:  $df=1$ ,  $F=8.87$ ,  $p=0.0029$ ). *Post hoc* weighted *t*-tests confirmed previous findings that static conspecific face cues elicited a significant congruence effect (RT difference: 5.3 ms, weighted two-tailed *t*-test:  $p=0.05$ ). A congruence effect was also observed for a simple off-center dot stimulus (RT difference: 7.9 ms,  $p<0.001$ ). In contrast, human- and cartoon-face cues did not elicit significant congruence effects. The overall reaction time distributions for congruent and incongruent trials are shown in **Figure 4B**.

In nature, gaze cues rarely occur as static images, but are signaled by a gesture, such as gaze shift or glance. We mimicked gaze shifts with a two-frame motion stimulus, which we refer to as a “dynamic cue.” In the first frame, the eyes of the cue face looked directly at the subject. In the second frame, the eyes were diverted



either to the left or the right. The time interval between the onset of the first and second frame was 17 ms. Dynamic face and non-face stimuli were used in the RO task to assess the impact of motion on the congruence effect (**Figure 4C**). A main congruence effect was observed (ANOVA, main effect:  $df = 1$ ,  $F = 72.38$ ,  $p < 0.001$ ). Motion enhanced the congruence effect for monkey-face cues (RT difference: 6 ms,  $p < 0.001$ ), non-social off-center dot cues (RT difference: 10.2 ms,  $p < 0.0001$ ), cartoon-face cues (RT difference: 5.5 ms,  $p = 0.0037$ ), and human-face cues (RT difference: 8.8 ms,  $p < 0.0001$ ). While the magnitude of the congruence effect was greatly enhanced for human-face cues (10 ms) and cartoon-face cues (4 ms), motion enhanced the congruence effect of conspecific face cues by a negligible amount (0.7 ms). The overall reaction time distributions for congruent and incongruent dynamic trials are shown in **Figure 4D**.

### LEARNED ORIENTING TASK

While the RO task examined the reflexive orienting of attention through face and non-face cues, the Learned Orienting (LO) task investigated the role of reward predictability on the orienting of attention. In the LO task, the gaze direction of the face stimuli and the spatial offset of non-face stimuli indicated which of two peripheral target was associated with reward. The numbers of sessions and trials for each monkey were as follows: Monkey C: 9 sessions, 5224 trials; Monkey F: 8 sessions, 5418 trials. During the learning period of the LO task, the accuracy of both subjects steadily improved, and the RT differences between correct and incorrect trials also increased with learning (**Figures 5A,B**). Additionally, both subjects demonstrated the biggest learning effect for non-face off-center dot cues, followed by monkey-face cues (**Figures 5C,D**).

After the training period, the LO task was modified to include 30% of trials with a single peripheral target (0 distractor trials). For these trials, the cue did not predict the location of the peripheral target (50% congruent-to-cue). The 0 distractor trials were thus identical to the dynamic cue version of the RO task, and serve as a cross-task comparison between reflexive orienting and learned orienting. It should be noted that these 0 distractor trials in the LO task were governed by somewhat different rules than the

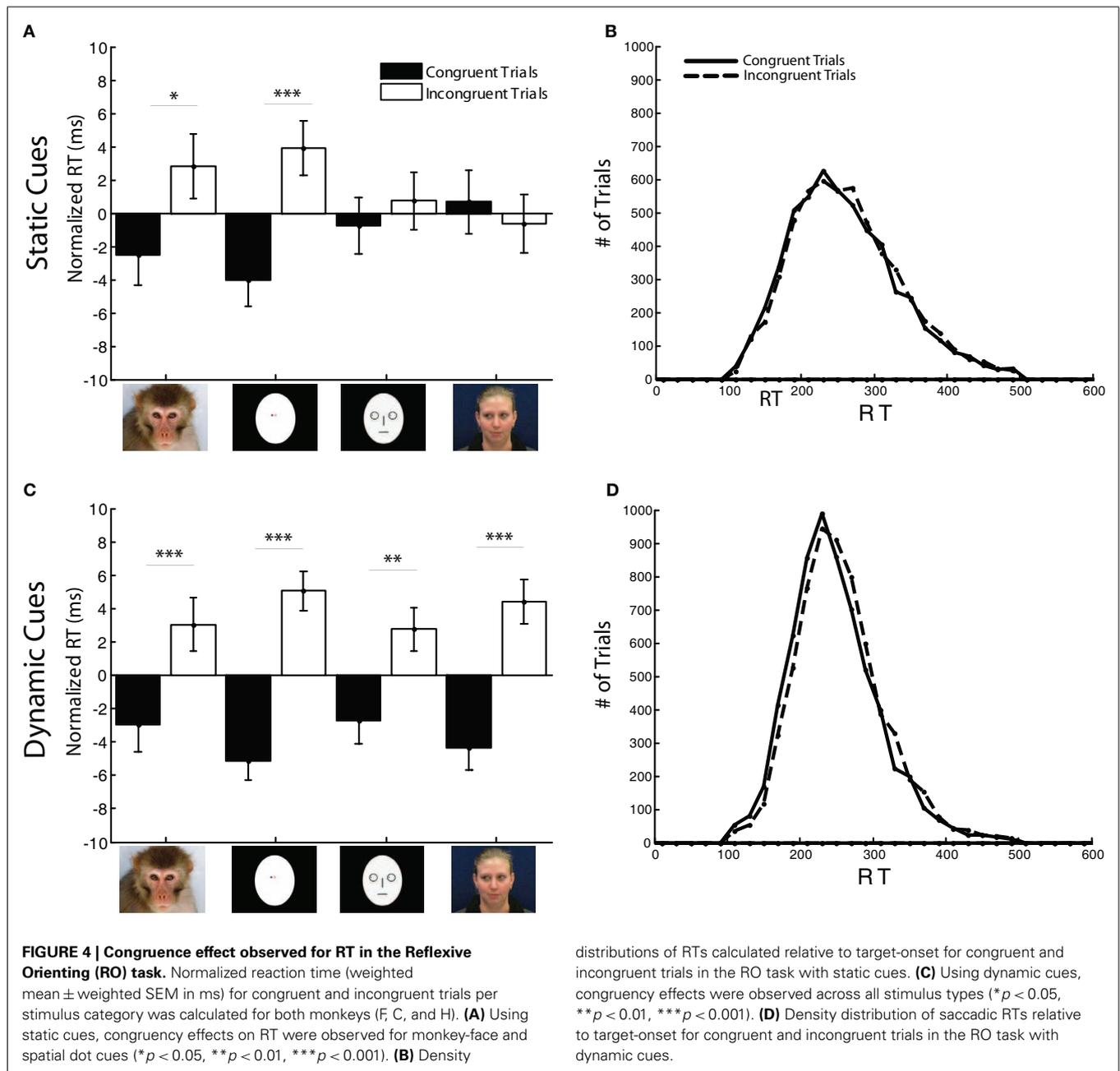
remaining 70% of trials, which had 1 distractor. In the 0 distractor trials, the rule was “look at the peripheral stimulus regardless of the direction of the cue.” In the 1 distractor trials, the rule was “look at the peripheral target that is congruent with the cue.”

The effect of congruence for the 0 distractor trials in the LO task was similar to that observed in the dynamic cue version of RO task (ANOVA, main effect:  $df = 1$ ,  $F = 31.32$ ,  $p < 0.001$ ; **Figure 6A**). Significant congruence effects were observed for monkey-face cues (weighted two-tailed  $t$ -test:  $p = 0.003$ ), non-face off-center dot cues ( $p = 0.0001$ ), and human-face cues ( $p = 0.014$ ), but not for cartoon-face cues.

Accuracy for 1 distractor trials (**Figure 6B**) was above chance but far short of ideal performance (mean percentage correct: monkey-face = 62.55, off-center dot = 67.89, cartoon-face = 56.08, human-face = 57.59). This could possibly be attributed to the partial conflict between 0 and 1 distractor trials as noted above. Mean RT for correct trials was significantly shorter than mean RT for incorrect trials (ANOVA, main effect on RT difference:  $df = 1$ ,  $F = 125.52$ ,  $p < 0.0001$ ). The magnitude of the RT difference was greatest for monkey-face cues, but significant for all stimulus categories (monkey-face: RT difference = 18.8 ms,  $p < 0.0001$ , off-center dot: 10.7 ms,  $p < 0.0001$ , cartoon-face: 12.5 ms,  $p < 0.0001$ , human-face: RT difference = 10.1 ms,  $p < 0.0001$ ).

### ANTICIPATORY SACCADDES

We defined “reactive” saccades as those initiated more than 100 ms after the onset of the peripheral target. In both the LO and RO task, over 90% of saccades made were reactive saccades. The remaining population of saccades, which we termed anticipatory saccades, had RT less than 100 ms (some of these saccades had negative RT, i.e., they were initiated prior to target-onset). Reactive saccades might be driven by both the cue and the peripheral target. Anticipatory saccades have only 100 ms (or less) to incorporate information about the peripheral target, but up to 550 ms to incorporate information about the cue. Hence, anticipatory saccades might reflect a tendency to imitate or “mirror” the behavior of the cue stimulus. In this case, one would expect the majority of anticipatory saccades to be congruent with the cued direction and

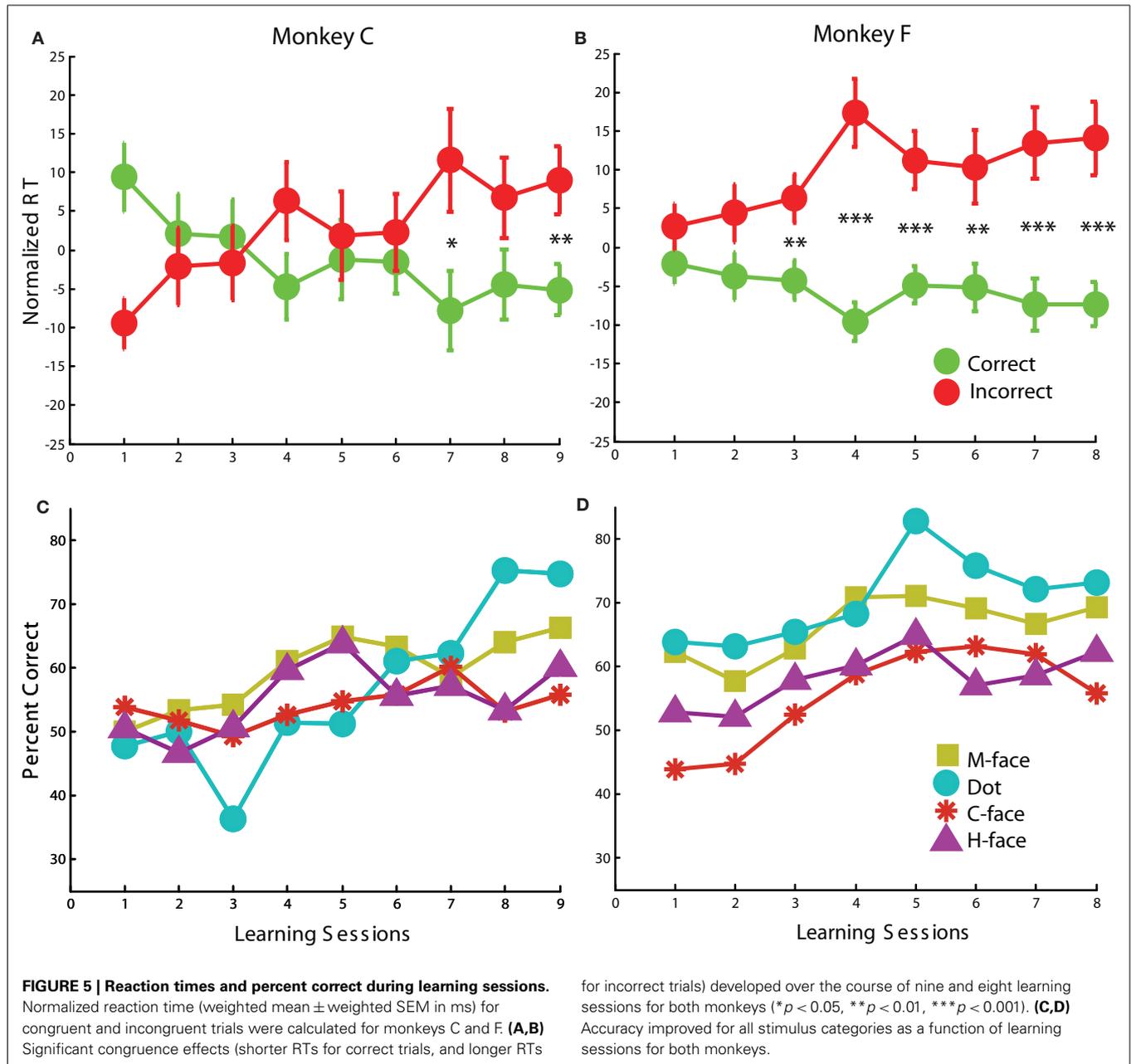


to occur more frequently for cues that elicit stronger automatic orienting.

In the RO task, the cues that evoked the highest proportion of anticipatory saccades were the monkey faces. The total number of anticipatory saccades was two to six times greater for monkey faces than for the other cue categories (**Figure 7A** static trials: 1.3%, **Figure 7B** dynamic trials: 5.4%). For static cues (**Figure 7A**), anticipatory saccades were more likely to mirror the cue if it was a monkey-face cue (binominal test:  $p = 0.049$ ) or a cartoon-face cue (binominal test:  $p = 0.037$ ). However, for dynamic cues, the anticipatory saccades did not show a significant tendency to mirror the cue stimulus; monkeys were equally likely to make an anticipatory saccade in the same direction as the cue (congruent) as they were

to saccade in the opposite direction (incongruent, **Figure 7B**). The total number of anticipatory saccades increased for dynamic monkey-face cues and cartoon-face cues.

In comparison to the RO task, the percentage of trials with anticipatory saccades was dramatically higher in the LO task (**Figure 7C** 0 distractor trials: 5.1%, **Figure 7D** 1 distractor trials: 7.3%). This is likely due to the fact that in the LO task, the cues predicted the rewarded target location. While this is strictly true only for 1 distractor trials, it has to be kept in mind that the number of distractors was unknown prior to the onset of the peripheral target. An increase in anticipatory saccades was observed in each stimulus category. Binominal tests showed that percentages of congruent-to-cue anticipatory saccades were above chance for all stimulus



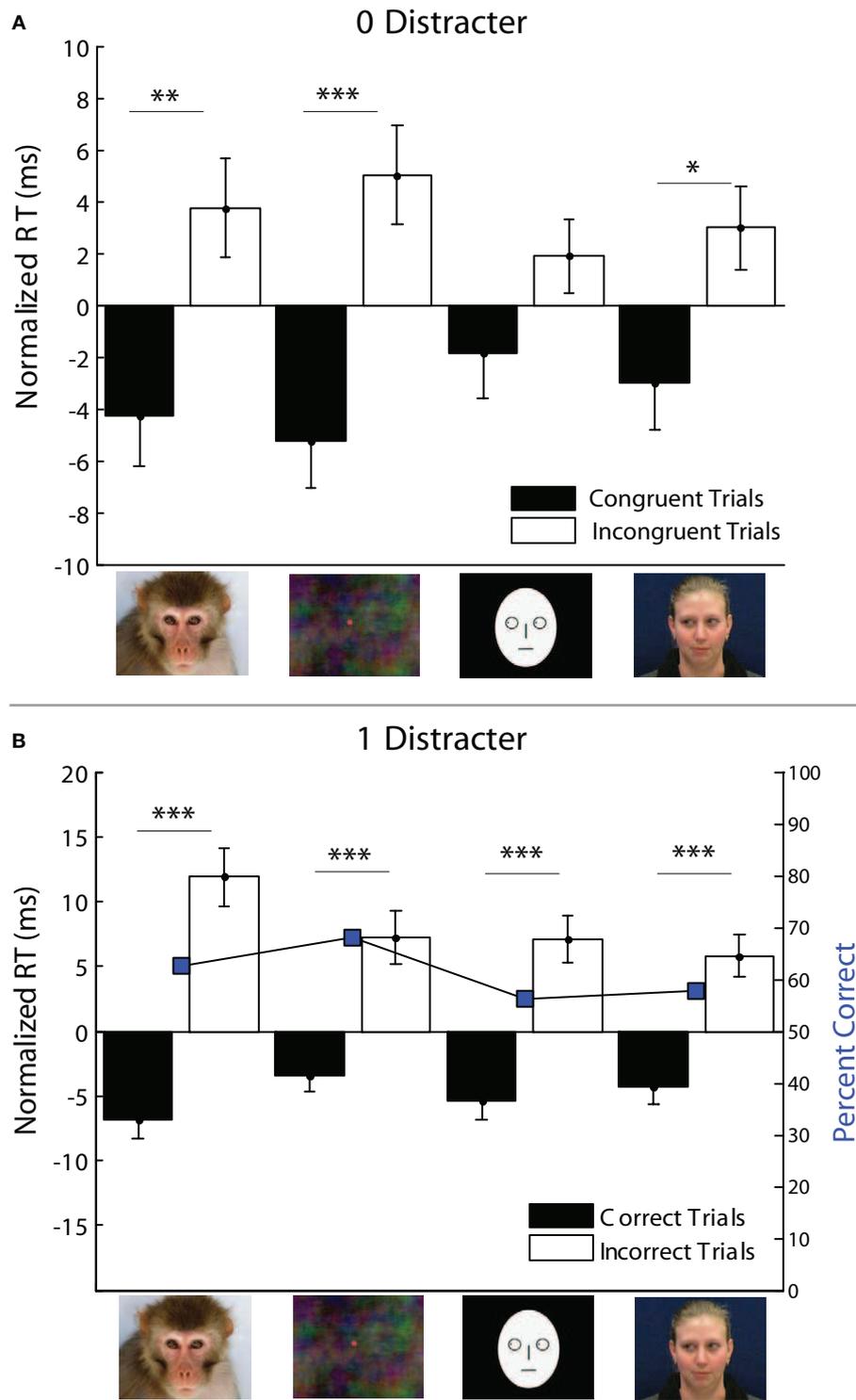
categories (0 distractor: monkey-face:  $p = 0.0015$ , off-center dot:  $p < 0.0001$ , cartoon-face:  $p < 0.0025$ , human-face:  $p < 0.0025$ ; 1 distractor: monkey-face:  $p < 0.0001$ , off-center dot:  $p < 0.0001$ , cartoon-face:  $p = 0.0002$ , human-face:  $p < 0.012$ ).

**MICROSACCADES**

Monkeys often made small saccades while viewing the cue stimulus. Such saccades were generally directed toward features of interest, such as the eyes. The frequency of such saccades could be considered an index of the degree to which the monkey was engaged or interested in the cue stimulus. To quantify the monkeys' level of interest in the cue stimuli, we classified rapid eye movements with amplitudes of less than  $1.0^\circ$  as "microsaccades"

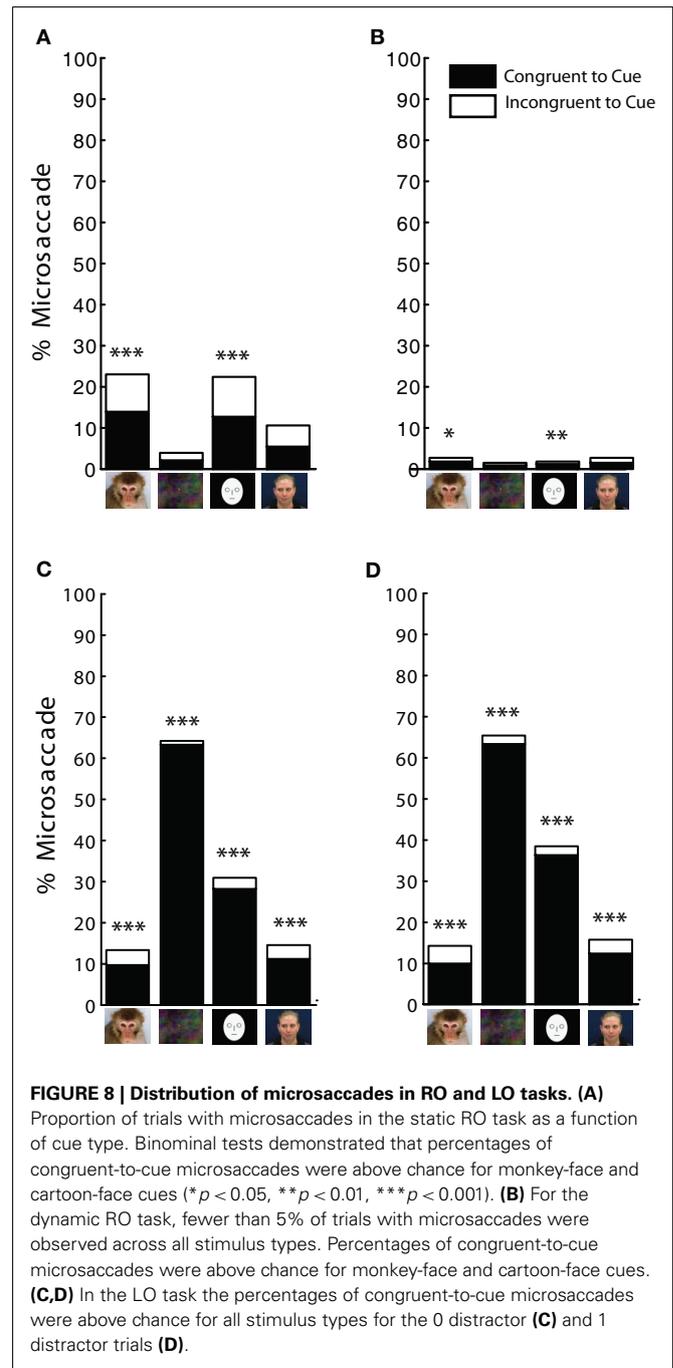
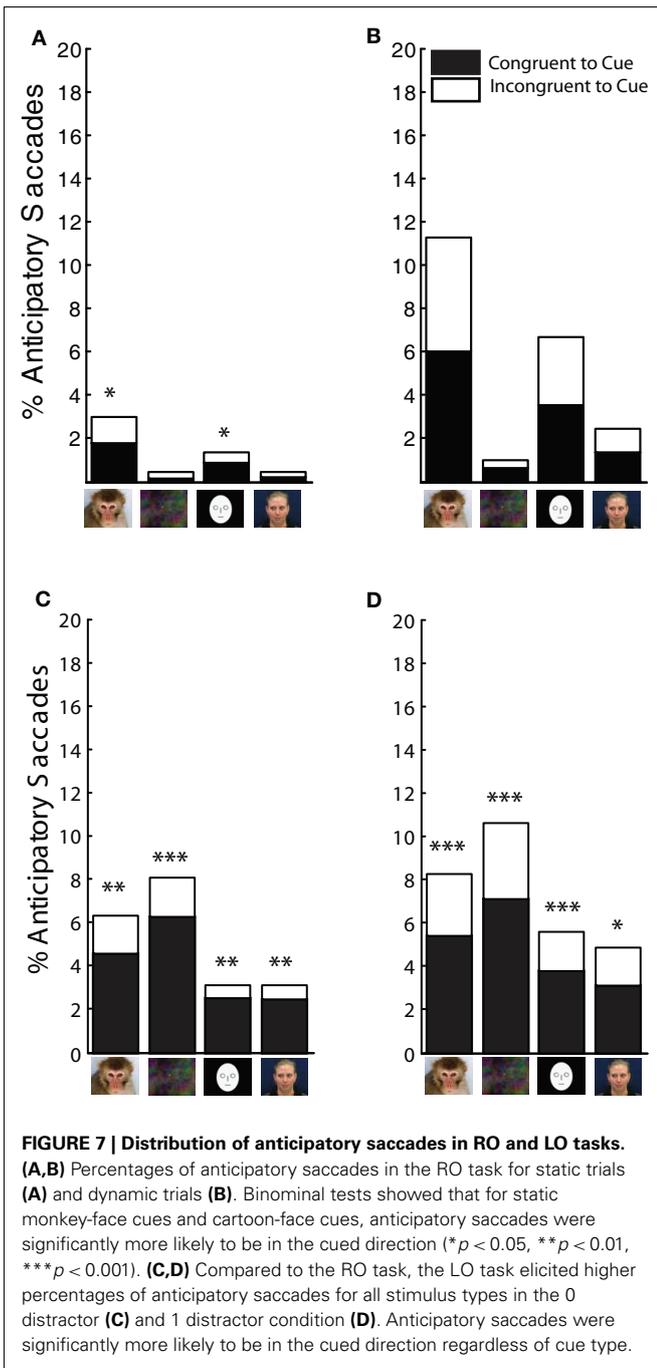
(Kliegl et al., 2009). We then calculated the proportion of trials with at least one microsaccade and divided these trials according to whether or not the microsaccade was congruent with the cue stimulus.

In the static cue version of the RO task, trials with monkey-face cues and cartoon-face cues had higher percentages of microsaccades (monkey-face = 23%, off-center dot = 3.7%, cartoon-face = 22.4%, human-face = 10.4%). The proportion of congruent-to-cue microsaccades (Figure 8A) was above chance for monkey-face cues and cartoon-face cues (Static cues: monkey-face:  $p < 0.001$ , off-center dot:  $p = 0.39$ , cartoon-face:  $p < 0.001$ , human-face:  $p = 0.26$ ). For the dynamic cue version of the RO task (Figure 8B), less than 5% of trials contained microsaccades for all



**FIGURE 6 | Congruence effect for social and non-social cues in the LO task. (A)** In the 0 distracter (single target) condition, congruence effects were observed for monkey-face, spatial dot, and human-face cues ( $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.001$ ). **(B)** In the 1 distracter condition, mean percent correct and weighted mean RT differences

between correct and incorrect trials were plotted. RT differences between correct and incorrect trials were significant across all stimulus types, with shorter latencies for congruent saccades. Accuracy was highest for off-center dots (68%) and followed by monkey-face cues (63%).



stimulus categories (percentages: monkey-face = 2.6%, off-center dot = 1.3%, cartoon-face = 1.6%, human-face = 2.5%). Despite the reduced number of microsaccades, binominal tests showed that percentages of congruent-to-cue microsaccades were above chance for monkey-face and cartoon-face cues (Dynamic cues: monkey-face:  $p = 0.0315$ , off-center dot:  $p = 0.0557$ , cartoon-face:  $p = 0.0088$ , human-face:  $p = 0.2601$ ).

The percentages of trials with microsaccades increased dramatically in the LO task (**Figures 8C,D**), presumably because the cues predicted reward location (percentages for

0 distractor trials: monkey-face = 13%, off-center dot = 63%, cartoon-face = 29%, human-face = 14%; percentages for 1 distractor trials: monkey-face = 14%, off-center dot = 63%, cartoon-face = 37%, human-face = 16%). The percentages of congruent-to-cue microsaccades were above chance for all stimulus categories (0 Distractor: monkey-face:  $p < 0.0001$ , off-center dot:  $p < 0.0001$ , cartoon-face:  $p < 0.0001$ , human-face:  $p < 0.0001$ ; 1 Distractor: monkey-face:  $p < 0.0001$ , off-center dot:  $p < 0.0001$ , cartoon-face:  $p < 0.0001$ , human-face:  $p < 0.0001$ ).

## DISCUSSION

The social orienting hypothesis posits a natural link between social cues such as gaze direction and visual spatial attention. Recent work suggests that there may be distinct neural substrates for social and non-social attention (Greene et al., 2009; Klein et al., 2009). This idea can be broken down into several testable sub-hypotheses: Social cues should direct attention automatically, without the need for training. Social cues should be more potent in directing attention than non-social cues with comparable physical properties. Cues from conspecifics should be more potent than cues from other species or artificial (cartoon) images. More realistic dynamic gestural cues should be more potent than static postural cues. We tested all of these predictions in a series of experiments in which we recorded the eye movements of three rhesus monkey subjects. We confirmed the advantage of conspecific faces over that of other species and cartoons, but not over non-face stimuli (off-center dots). We found that adding motion in the form of a simulated glance had very little to no effect on enhancing the congruence effect for conspecific faces, but had a significant effect for human faces. When the cues predicted reward, subjects learned the association between cue direction and reward location. After learning, the accuracy and choice reaction time increased for all stimuli, particularly for conspecific faces. However, the congruence effect on reaction time remained largest for non-face stimuli.

### AUTOMATICITY AND SPECIFICITY

We confirmed previous reports (Deaner and Platt, 2003) that gaze cues (pictures of conspecifics faces looking to the left or right) are associated with reflexive orienting of attention in rhesus macaques (Figure 4A). We used a Posner paradigm in which the cued direction was not predictive of the location of the peripheral saccade target. Nevertheless, response times for saccades to the peripheral target were faster when the target location was congruent with the gaze direction of the conspecific (monkey) face cue and slower when the target and cue were incongruent. The magnitude of the congruence effect on saccade initiation for conspecific face cues averaged about 5.3 ms. Due to our small sample size, we also tested for inter-subject variability and we did not find a significant effect across subjects. It is important to keep in mind that our observations were not conducted in the wild but in a captive population. Since these monkeys have been living in confinement since birth and with limited contact from other monkeys as adults, we did not address the issue of dominance in this study. All pictures of monkeys used in this study were unfamiliar monkeys with neutral expressions chosen from the database of pictures from Yerkes Regional Primate Center. Monkeys had been previously trained to perform visual psychophysical and oculomotor tasks such as discriminating the speed of random dot patterns or tracking moving targets. The random dot stimuli were quite different than the cue stimuli used in the current study, and it is difficult to imagine that there would be any carryover from discriminating random dot motion to attending to faces or single dots. Tracking moving targets is something that monkeys do naturally (indeed, they can often be “trained” to perform smooth pursuit tasks in a matter of minutes).

The congruence effect for gaze cues is typically interpreted as resulting from a complex series of neural processes: (1) animals

decode information about the direction of gaze of conspecifics from a complex visual stimulus, (2) they interpret the implicit social message, i.e., “something interesting is happening over there,” and (3) this message automatically directs their attention to the cued location. However, so far, a simpler alternative has not been excluded: the attention cue is encoded by means of a physical feature, i.e., the relative positions of the pupils in the eyes. It is in principle possible that this physical feature alone directs attention in a manner similar to a non-social spatial cue, in our case the off-center dot. Our data allow us to dissociate the two mechanisms. In particular, we show that human and cartoon faces that share a physical feature (relative position of pupils in the eyes) do not appear to guide attention. Only monkey faces reliably guided attention toward the cued direction. One would like to conclude that the effect of monkey-face cues is due to the social message they encode. However, this interpretation is clouded by the fact that a simple offset dot cue produced effects comparable to the monkey faces. This control condition was not present in Deaner and Platt (2003), presumably because no effect was expected. The finding of a robust effect raises the possibility that purely abstract stimuli can produce effects comparable to social stimuli. Thus, while one can conclude that attending to monkey faces conferred an attentional advantage over human and cartoon faces (and this may be due to social attention), it was not the case that monkey faces were the most potent attention cues overall.

### POSTURAL VS. GESTURAL CUES

Gestures may be more realistic and salient than static postural cues, and hence more effective in directing attention. Gestures may also induce a “mirroring” response in which the subject imitates the movement performed by the cue image. We tested this by designing cue stimuli that simulated a “glance” with two-frame motion. We found that this dynamic gestural cue enhanced the congruence effect by at least 2 ms for all stimuli except monkey-face cues (Figure 4C). For monkey faces, the effect of adding motion was only 0.7 ms more than the static cue (Figure 4A). Adding motion dramatically increased the percentage of anticipatory saccades for all stimuli (compare Figure 7A, static, vs. Figure 7B, dynamic), but these anticipatory saccades were equally likely to be congruent or incongruent with the direction of motion in the cue.

One possible explanation for the weak effects of motion in the monkey-face cues is the fact that, unlike a human eye, the rhesus eye has low contrast between the pupil, iris, and sclera, thus creating the least amount of motion energy (Figure 2B). It has been suggested that the human eye is uniquely suited to promote fast discrimination of gaze direction, for it has a much higher dark-iris-to-white contrast compared to non-human primate eyes (Kobayashi and Koshima, 1997). We computed motion energy (Figure 2B) for each image class individually to assess whether or not it was a significant factor in determining congruence effects. We found that cartoon faces and dots possessed the strongest motion energy due to the fact that their features had the highest contrast. However, motion in the cue stimulus had the biggest behavioral effect for human faces, despite their comparatively low motion energy. Hence, it seems unlikely that motion energy was a significant factor in orienting of attention. In fact, while static human-face cues failed to elicit a significant effect, dynamic

human-face cues did produce a significant congruence effect. This suggests that static human faces lack the intrinsic value that reflexively drives social attention. However, in the context of a dynamic gesture, which is more realistic, human faces may come to acquire a similar degree of social significance as conspecific faces.

Because cartoon faces were made up of high contrast edges, these images had the largest Fourier amplitude (Figure 2A) across all spatial frequencies and also the highest motion energy (Figure 2B). However, both static and dynamic cartoon faces were consistently the least effective stimuli for guiding attention. We can infer that high contrast image features, such as the contour lines that define a cartoon face, are not sufficient to drive attention, thus explaining why cartoon faces may not function effectively as social cues.

## EFFECTS OF LEARNING

The current findings raise a key question: are macaques able to decode the social message inherent in human and cartoon face images? If so, the lack of an effect for these stimuli would arise because monkeys do not act on the social message, presumably because they do not consider the focus of attention of a human (or cartoon) subject worthwhile. This could be changed if the monkeys learn that the direction of gaze of human or cartoon faces is predictive of reward. Thus, in addition to reflexive orienting of attention, we also studied overt choice behavior in response to face and non-face cues in the Learned Orienting (LO) task. We made both the face and non-face stimuli equally valuable in terms of predicting the rewarded target location.

Accuracy (percent correct) for the LO task improved over sessions for all stimulus categories (Figures 5C,D). However, subjects demonstrated the best performance for non-face off-center dot cues. Performance for monkey faces was better than for human and cartoon faces. Thus, monkey faces yielded better performance than other face stimuli, but not the best accuracy overall.

After learning, there was an increase in the choice reaction time difference between correct and incorrect trials averaged over all stimuli (Figures 5A,B). This is expected if learning influences a diffusion-to-bound decision process (Ratcliff and McKoon, 2008) by increasing the rate of evidence accumulation for the cued target. While the choice RT difference was significant for all stimuli, the effect was roughly twice as large for monkey faces than for other stimuli (Figure 6B). Taken together, accuracy and RT data suggest that monkey faces were most effective in inducing learning.

## REFERENCES

- Adelson, E. H., and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A* 2, 284–299.
- Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.
- Deaner, R. O., and Platt, M. L. (2003). Reflexive social attention in monkeys and humans. *Curr. Biol.* 13, 1609–1613.
- Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., and Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Vis. Cogn.* 6, 509–540.
- Friesen, C. K., and Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychon. Bull. Rev.* 5, 490–495.
- Friesen, C. K., and Kingstone, A. (2003). Covert and overt orienting to gaze direction cues and the effects of fixation offset. *Neuroreport* 14, 489–493.
- Friesen, C. L., Ristic, J., and Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 319–329.
- Greene, D. J., Mooshagian, E., Kaplant, J. T., Zaidel, E., and Iacoboni, M. (2009). The neural correlates of social attention: automatic orienting to social and nonsocial cues. *Psychol. Res.* 73, 499–511.
- Hill, J. L., Patel, S., Gu, X., Seyedali, N. S., Bachevalier, J., and Sereno, A. B. (2010). Social orienting: reflexive versus voluntary control. *Vision Res.* 50, 2080–2092.
- Horner, V., Carter, J. D., Suchak, M., and de Waal, F. B. M. (2011). Spontaneous prosocial choice by chimpanzees. *Proc. Natl. Acad. Sci.* 108, 13847–13851.
- Jonides, J. (1981). “Voluntary versus automatic control over the mind’s eye’s movement,” in *Attention and Performance*, Vol. 9, eds J. B. Long and A. D. Baddeley (Hillsdale, NJ: Erlbaum), 187–203.
- Klein, J. T., Shepherd, S. V., and Platt, M. L. (2009). Social attention and the brain. *Curr. Biol.* 19, R958–R962.

For trials with no distractor, learning did not increase the congruence effect on reaction time for any stimulus category (compare Figures 4C and 6A). However, learning did increase the overall frequency of anticipatory saccades and microsaccades as well as the proportion of these saccades that were congruent with the cue. These effects were strongest for non-face cues.

To interpret results from the learned orienting task, we must consider both the decision-making and attention components of the task. The aspects of the task that reflect decision-making (accuracy and choice RT) were enhanced by learning and this enhancement was strongest for monkey faces and non-face cues. Given that accuracy was about the same for both monkey faces and non-face cues, but that choice RT differences were much greater for monkey faces, the latter cue category appears to mediate the strongest effects of learning on decision-making.

Aspects of the task that are more closely associated with attention (congruence effect, anticipatory saccades, microsaccades) were either not enhanced by learning or enhanced most for non-face stimuli. For face stimuli, the effects of learning appear to be stronger for behavioral outputs that are associated with decision-making than those associated with attention.

## GENERAL CONCLUSIONS

Taken as a whole, the findings of the current study suggest that gaze direction cues, which are sometimes assumed to be inherently social, can direct attention automatically regardless of the reward value associated with the cue. This effect is species-specific, as attentional effects for monkey faces were always stronger than those for human and cartoon faces. However, all face stimuli showed weaker effects when compared to the most rudimentary non-social stimuli. This study also showed that motion generally enhances the magnitude and reliability of attentional effects for all stimuli. When the cues were predictive of reward, monkeys learned this association for all stimuli, regardless of social or non-social content. Learning affected both attention-related and decision-related aspects of the task. The decision-related learning was strongest for monkey faces, while attention-related learning was strongest for non-face stimuli.

## ACKNOWLEDGMENTS

This work was supported by MH059244 (Vincent P. Ferrera) and DFG (Tobias Teichert). The authors wish to thank Joyce Cohen, Jeff Fisher, and Mark Wilson for pictures of monkeys from the Yerkes Regional Primate Center.

- Kliegl, R., Rolfs, M., Laubrock, J., and Engbert, R. (2009). Microsaccadic modulation of response times in spatial attention tasks. *Psychol. Res.* 73, 136–146.
- Kobayashi, H., and Koshima, S. (1997). Unique morphology of the human eye. *Nature* 387, 767–768.
- Langton, S. R. (2000). The mutual influence of gaze and head orientation in the analysis of social attention direction. *Q. J. Exp. Psychol. A* 53, 825–845.
- Langton, S. R. H., and Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Vis. Cogn.* 6, 541–567.
- Martin, C. F., Biro, D., and Matsuzawa, T. (2011). Chimpanzees' use of conspecific cues in matching-to-sample tasks: public information use in a fully automated testing environment. *Anim. Cogn.* 14, 893–902.
- Nakayama, K., and Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Res.* 29, 1631–1647.
- Nummenmaa, L., and Calder, A. J. (2009). Neural mechanisms of social attention. *Trends Cogn. Sci.* 13, 135–143.
- Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25.
- Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922.
- Ristic, J., and Kingstone, A. (2005). Taking control of reflexive social attention. *Cognition* 94, B55–B65.
- Ristic, J., Wright, A., and Kingstone, A. (2007). Attentional control and reflexive orienting to gaze and arrow cues. *Psychon. Bull. Rev.* 14, 964–969.
- Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces. *Proc. Natl. Acad. Sci. U.S.A.* 105, 394–398.
- Sanderson, C., and Paliwal, K. K. (2002). “Polynomial features for robust face authentication,” in *IEEE International Conference on Image Processing (ICIP)*, Rochester.
- Teufel, C., Gutmann, A., Pirow, R., and Fischer, J. (2010). Facial expressions modulate the ontogenetic trajectory of gaze-following among monkeys. *Dev. Sci.* 13, 913–922.
- Tipples, J. (2008). Orienting to counterpredictive gaze and arrow cues. *Percept. Psychophys.* 70, 77–87.
- Tomonaga, M., Tanaka, M., Matsuzawa, T., Myowa-Yamakoshi, M., Kosugi, D., Mizuno, Y., Okamoto, S., Yamaguchi, M. K., and Bard, K. A. (2004). Development of social cognition in infant chimpanzees (*Pan troglodytes*): face recognition, smiling, gaze, and the lack of triadic interactions. *Jpn. Psychol. Sci.* 46, 227–235.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 19 September 2011; accepted: 30 May 2012; published online: 25 June 2012.

Citation: Yu D, Teichert T and Ferrera VP (2012) Orienting of attention to gaze direction cues in rhesus macaques: species-specificity, and effects of cue motion and reward predictiveness. *Front. Psychology* 3:202. doi: 10.3389/fpsyg.2012.00202

This article was submitted to *Frontiers in Comparative Psychology*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 Yu, Teichert and Ferrera. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.