PROCEEDINGS B

Research

CrossMark
click for updates

Author for correspondence:
Karolina Marciniak
e-mail: marciniak.kar@gmail.com

THE ROYAL SOCIETY
PUBLISHING

Monkeys head-gaze following is fast, precise and not fully suppressible

Karolina Marciniak, Peter W. Dicke and Peter Thier

Department of Cognitive Neurology, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

Human eye-gaze is a powerful stimulus, drawing the observer's attention to places and objects of interest to someone else ('eye-gaze following'). The largely homogeneous eyes of monkeys, compromising the assessment of eye-gaze by conspecifics from larger distances, explain the absence of comparable eye-gaze following in these animals. Yet, monkeys are able to use peer head orientation to shift attention ('head-gaze following'). How similar are monkeys' head-gaze and human eye-gaze following? To address this question, we trained rhesus monkeys to make saccades to targets, either identified by the head-gaze of demonstrator monkeys or, alternatively, identified by learned associations between the demonstrators' facial identities and the targets (gaze versus identity following). In a variant of this task that occurred at random, the instruction to follow head-gaze or identity was replaced in the course of a trial by the new rule to detect a change of luminance of one of the saccade targets. Although this change-of-rule rendered the demonstrator portraits irrelevant, they nevertheless influenced performance, reflecting a precise redistribution of spatial attention. The specific features depended on whether the initial rule was head-gaze or identity following: head-gaze caused an insuppressible shift of attention to the target gazed at by the demonstrator, whereas identity matching prompted much later shifts of attention, however, only if the initial rule had been identity following. Furthermore, shifts of attention prompted by head-gaze were spatially precise. Automaticity and swiftness, spatial precision and limited executive control characterizing monkeys' head-gaze following are key features of human eye-gaze following. This similarity supports the notion that both may rely on the same conserved neural circuitry.

1. Introduction

Successful social interactions require understanding peer dispositions, desires, beliefs and intentions. A major step in developing this theory of (others') mind is the ability to shift attention to the same location and/or object another person is interested in, i.e. to establish joint attention [1].

For the observer, the direction of another person's eyes is a major source of information on the object or place of interest to that person. Human observers experience a strong urge to follow peer eye-gaze either overtly, by making an eye movement themselves, or by shifting attention covertly. These shifts of attention cannot be suppressed by a primary interest in some other place or object [2] and not even by prior knowledge that the other's gaze may actually be misleading [3]. These observations suggest that human eye-gaze following is a largely automatic or reflex-like behaviour akin to the one evoked by salient, sudden-onset peripheral ('exogenous') stimuli [4]. In standard spatial cueing paradigms, using such exogenous stimuli, subjects respond faster to the cued than to the non-cued target when the cue-target interval is short ('response facilitation'). This pattern gets inverted for longer cue-target intervals ('inhibition of return', IOR) [5]. Eye-gaze cues elicit very similar orienting effects as measured by reaction times as standard exogenous cues. However, there are also subtle differences between them with respect to the maintenance and quality of the cueing effects across time: eye-gaze cueing causes longer response facilitation compared to standard exogenous cues and an IOR occurs for eye-gaze following only at cue-target intervals that are much longer than the ones for exogenous cues [6,7]. In other words, the observer seems to be reluctant to withdraw attention

from the gazed-at location [6], arguably because the attention-binding effect of gaze is particularly powerful. In sum, human eye-gaze following is a socially relevant automatic process, most probably relying on preformed brain circuitry. This notion receives support from developmental studies showing that newborns already exhibit sensitivity to eye-like stimuli driving attention [8]; and at three months of age, infants use the other's eye-gaze to shift their attention to peripheral probes [9]. Recent lesion and human functional magnetic resonance imaging (fMRI) studies suggest that a region in the posterior superior temporal sulcus is the core of the brain circuitry supporting this domain-specific process [10,11].

Also adult monkeys can shift their attention guided by eye-gaze, provided the eyes are seen from close distance [12–14]. However, the lack of conspicuous features of a monkey's eye [15] prevent the usage of eye-gaze cues at least from larger distances and may in general decrease the relevance of this cue relative to the orientation of the conspecific's head, directed at objects of interest, i.e. head-gaze. Actually also human spatial attention may be influenced by a number of other bodily cues such as head or trunk orientation or pointing gestures. Yet, clearly the dominating cue is eye-gaze [16]. Given the fact that in monkeys, head-gaze seems to have the particular importance eye-gaze has for humans, we asked if monkeys' head-gaze following may exhibit similar functional characteristics as human eye-gaze following? Previous work seemed to suggest that both the time course and the attentional benefit of monkeys' head-gaze following and human eye-gaze following may be similar [12]. However, it is unclear if monkeys' head-gaze following also exhibits the relative independence from high-level cognitive processes that underlie the automaticity of human eye-gaze following [2]. So far, the evidence for automaticity of monkeys' head-gaze following is largely circumstantial and, in any case, insufficient to shed light on the detailed structure of the behaviour. The reason is that in previous studies of monkeys' head-gaze following, the tasks did not involve elements in which the head-gaze cue would not only have been task-irrelevant but actually misdirecting and therefore requiring cognitive control. In order to clarify if monkeys' head-gaze following can be controlled if required and, moreover, to characterize the time course of head-gaze following, we trained rhesus monkeys both on a head-gaze following task as well as on a second task, requiring the usage of a learned association between particular spatial targets and distinct facial identities (head-gaze following versus identity matching) in the latter case ignoring the possibly conflicting information provided by head orientation. The attentional shifts evoked by these tasks were gauged by asking the monkeys to detect a luminance change at distinct spatial positions. The relevant location could correspond to the location determined by head-gaze, by facial identity or be different from either of the two, independent of the prevailing task rule requiring head-gaze following or facial identity matching. We demonstrate that head-gaze cues cause early covert shifts of attention that cannot be suppressed, even if inappropriate as the prevailing task rule demanded ignoring them.

2. Material and methods

(a) Subjects
Two rhesus monkeys (*Macaca mulatta*): M1 (8 years, 9 kg); M2 (10 years, 12 kg) were implanted with three cf-PEEK

(carbon-fibre-enforced polyetheretherketone) tripods, each attached to the skull with six ceramic screws (Thomas Recording, Marburg, Germany). Surgeries were carried out under combination anaesthesia with isoflurane and remifentanyl with monitoring of all relevant vital parameters (body temperature, CO_2, blood oxygen saturation, blood pressure, ECG). After surgery, monkeys were supplied with opioid analgetics (buprenorphin) until full recovery. All animal preparations and procedures fully complied with the NIH Guide for Care and Use of Laboratory Animals, and were approved by the local animal care committee (RP Tübingen, FG Tierschutz).

(b) Training
The monkeys were trained in a horizontal primate chair for later studies in an MRI scanner. Their heads were fixed to the chair by screwing the tripods to an acrylic cap with an integrated massive cf-PEEK rod connected to the chair's frame. Eye position was tracked in real time using a low-cost CMOS-infrared camera (C-MOS-Kameramodul1(C-CAM-A), Conrad Elektronik, Germany) with infrared emitting LEDs. The custom-made software running on a standard PC determined the location of the centre of the pupil with a spatial resolution of 0.5° visual angle and a temporal resolution of 50 Hz.

(c) Experimental paradigms
(i) Experiment A: learning to use different types of cues for the guidance of spatial attention
Monkeys saw portraits of demonstrator monkeys on a monitor in front of them. They were extensively trained (M2: six months; M1: 1 year) to use either the orientation of the demonstrator monkey's head or the monkey's facial identity, associated with distinct spatial positions, to shift spatial attention. The features of the central fixation target determined the rule: a red circular fixation spot cued the observer to make a saccade to the demonstrator's gaze target (gaze following). In the case of a green rectangle, the saccade target was identified by learned associations between the peripheral target and the demonstrator's facial identity (identity matching; electronic supplementary material, figure S1). The behavioural results obtained with this paradigm have already been published in a report on the cortical underpinnings of gaze following in monkeys [17] and further details on the paradigm can be found there. At the centre of this study are experiments B and C which build on this initial experiment A.

(ii) Experiment B: assessing the role of stimulus duration
We used the above-mentioned basic paradigm to assess how the ability to shift spatial attention based on gaze direction or identity depended on the lifetime of the portraits ('stimulus duration'). To this end, we pseudo-randomly presented blocks of 20–30 trials of either gaze following or identity matching with varying stimulus duration (electronic supplementary material, figure S1c). Note that in this experiment, the demonstrator portrait was not preceded by the fixation portrait as in A.

(iii) Experiment C: probing the time course of spatial attention
Our primary aim was to probe the time course of monkeys' shifts of spatial attention guided by head-gaze or facial identity. This was achieved by determining the monkeys' ability to detect subtle transient changes in luminance levels at distinct spatial and temporal positions. In this experiment, a trial (figure 1) started with the central cue chosen randomly to indicate that either gaze following or identity matching would be required. Four hundred milliseconds later, the demonstrator portrait was presented for 300 or 400 ms (stimulus duration based on results from experiment B). One hundred or 200 ms later (depending on
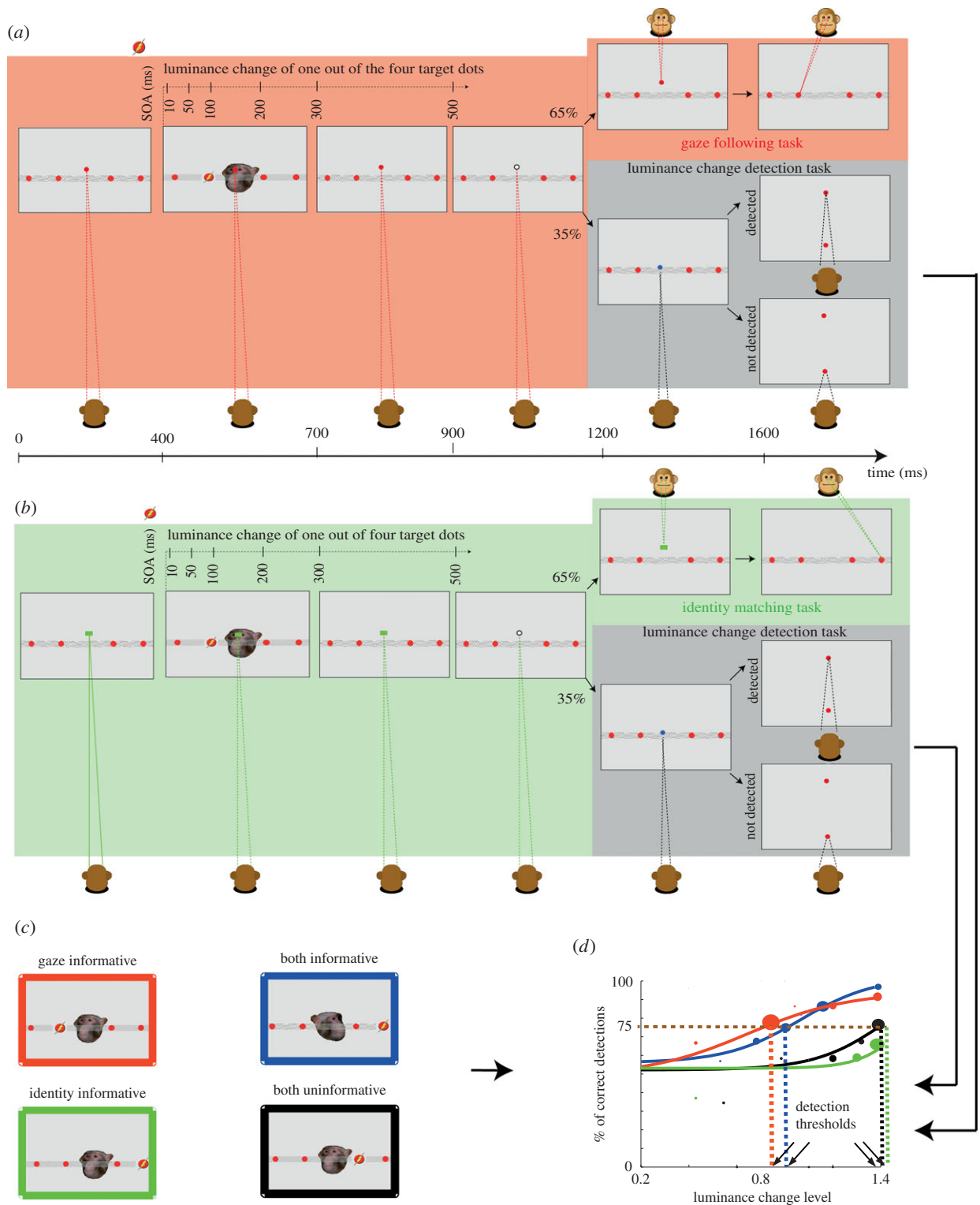
Figure 1. Paradigm for probing the time course of attentional shifts guided by head-gaze or facial identity. Sequence of events in experiment C in a gaze following (a) or identity matching context (b). The differently coloured backgrounds indicate the rule prevailing at a certain time: the gaze following (red), identity matching (green) rule, luminance detection (grey). (c) Cartoons, describing four congruency categories defined by the spatial relationship of the luminance change target with the target cued by the portrait's head-gaze and its facial identity: gaze informative (red frame), identity informative (green frame), both informative (blue frame) and both uninformative (black frame). (d) Exemplary psychometric curves: the percentages of correct luminance change detections as a function of the luminance change level, for the four cases described in (c), fitted by logistic functions. At detection thresholds, the fits predicted 75% of correct detections.

the duration of the portrait: 400 or 300 ms), a central, neutral cue (a white circle) appeared for 300 ms, which was then replaced by an informative cue. In 65% of all trials, it was the initial cue for gaze following (red spot) or identity matching (green rectangle) ('standard trials'). In 35% of the trials, the informative cue adopted a new feature (blue spot) indicating a switch to a new task rule ('detection trials') demanding the detection of a transient (80 ms duration) change of luminance of variable degree, in 50% of all detection trials affecting one of the four peripheral targets. This luminance change took place in a period starting at portrait onset up to 500 ms later (stimulus-onset asynchrony relative to portrait onset, SOA). Note that changes in luminance ended well before providing the ultimately effective instruction, at a time the monkeys had to assume that they would probably be rewarded for gaze following or identity matching as called for by the initially presented central cue. The final disappearance of the instructive cue invited

4

the observer to deliver his response. Monkeys were rewarded for correctly identifying the spatial target as indicated by gaze or identity information in standard trials or by correctly detecting the presence or absence of a change in luminance in detection trials. Targets had a baseline luminance of 5 cd m^{-2} and the parameter estimation for sequential testing (PEST) strategy [18] determined and increased luminance that could reach a maximum of 43 cd m^{-2}. Luminance increases were described by their percentage values (baseline = 48%, maximal = 100%). Luminance change detection trials fell into four categories defined by the congruency of the spatial positions of the luminance change target with the positions singled out by the portrait's head-gaze and identity, respectively: both uninformative (neither gaze nor identity point to the luminance change target), both informative (both gaze and identity point to the luminance change target), identity informative (only identity points to the luminance change target) and gaze informative (only gaze points to the luminance change target).

We ran separate, pseudorandomized experimental sessions (involving about 450 trials each) for each SOA randomizing trial types within sessions (see above), collecting around 20 detection trials per category and initial rule in each session. The size of the luminance change was chosen by an adaptive staircase procedure (PEST strategy; [18]). We collected 112 experimental sessions in M1 and 130 experimental sessions in M2.

(d) Data analysis

To analyse the effect of stimulus duration on performance (experiment B), we ran five to six sessions per duration, condition (i.e. gaze following versus identity matching) and monkey, and calculated the mean performance across sessions. To compare the means, we deployed a one-way ANOVA with the factor portrait duration, separately for each monkey (a Kolomogov–Smirnoff test had confirmed normal distribution of the data).

In experiment C, we measured the sensitivity for luminance changes in each monkey, given SOA and category (each obtained from 14 to 18 experimental sessions) by fitting a logistic function to the plot of the percentage of correct responses as a function of the luminance change level. The luminance changes for which the fit predicted 75% of correct responses were taken as proxy of luminance change sensitivity ('detection threshold'). We had to discard 32% of all sessions in M1 and 26% in M2 as the logistic fits obtained were not significant (discussed later).

In order to fit the data with logistic functions, we used the routines of the Palamedes toolbox for MATLAB [19]. Because of the relatively small sample size available for individual condition, we preferred logistic fits to probit fits as the latter have been shown to be potentially inaccurate for smaller numbers of trials [20,21]. The logistic function involves four parameters: the guessing rate γ (probability that the guess would be correct), the lapse rate λ (probability of responding incorrectly as a result of a lapse), the detection threshold α (the stimulus value at 75% of correct detection responses) and the slope of the function at the detection threshold β. We assumed that γ and λ reflecting the general experimental conditions would be constant across category trials, whereas α and β would reflect the differences in detection performance, a proxy for the attentional modulation across categories. Therefore, keeping $\gamma = 0.5$ (the guessing rate = 50%) and $\lambda = 0$ (assuming subject's responses according to the luminance level), we searched for the best estimates of α and β. Moreover, for each logistic function, we calculated the goodness of fit (transformed likelihood ratio, Dev) and the associated p-value (pDev; range from 0 to 1, the larger the value the better the fit). Fits were considered significant and the resulting detection thresholds used if $p < 0.05$. All resulting parameters were obtained separately for each monkey, both for gaze following (electronic supplementary material, table S1) and identity matching (electronic supplementary material, table S2). To compare detection thresholds between conditions, we created a sampling distribution

for each one, using the bootstrap procedure of the Palamedes toolbox. The resulting threshold's sampling values were normalized across monkeys and across SOAs (both to the mean threshold) and used for further statistical analysis. Statistical comparison of detection thresholds was carried out by a $2 \times 6 \times 4$ repeated measures ANOVA with the factors condition (gaze following versus identity matching), SOA (10, 50, 100, 200, 300 and 500 ms) and category (gaze informative, identity informative, both informative and both uninformative). The sampling distributions were normally distributed (Kolmogorov–Smirnov test).

To assess the precision of the early gaze following effect, we used the above data and grouped it into the three new categories, based on spatial vicinity of the luminance change target to the gazed-at target: 'precise' (the luminance change occurs precisely at the gazed-at target), 'nearby' (the luminance change occurs at the target just adjacent to the gazed-at target within the same hemifield) and 'opposite' (the luminance change occurs at the target laying in the hemifield opposite to the gazed-at target). We repeated the data analysis procedure as described above for each monkey, each condition and two SOAs (50 and 100 ms). Statistical comparison of detection thresholds was carried out by resorting to a $2 \times 2 \times 3$ repeated measures ANOVA with the factors condition (gaze following and identity matching), SOA (50 and 100 ms) and precision (precise, nearby and opposite), the sampled distributions were normally distributed (Kolmogorov–Smirnov test).

3. Results

For this study, we used monkeys which had already been extensively trained to make accurate saccades towards distinct spatial targets, identified by head-gaze of the portrayed monkey or, alternatively, by relying on the learned association between particular targets and distinct facial identities (experiment A). As reported previously [17], they successfully learned the two tasks and their performance levels were very similar. Moreover, by running two behavioural control experiments, we could establish that the two monkeys indeed used head-gaze information in a geometrical manner in the gaze following task, rather than trying to exploit eventually learned associations between particular head orientations and targets. Experiment C was designed to gauge the time course and controllability of shifts of spatial attention prompted by gaze cues. In experiment C, in which a demonstrator monkey provided the cue that would reallocate attention, it was necessary to keep the presentation of the demonstrator as short as possible. This was necessary to accommodate SOAs between cue onset and the onset of the discriminandum, short enough to detect early shifts of attention. The purpose of experiment B was to determine the minimal presentation duration of the portrayed monkey (i.e. minimal cue duration) still evoking reliable shifts of attention. We assessed the role of cue duration in providing valid spatial information, either based on gaze direction or the identity of the portrayed monkey. For the longest stimulus duration, the two monkeys reached performance levels around 70% on both tasks and tended to deteriorate gradually and monotonically with decreasing duration, falling below the 50% level at durations of 200–300 ms. At durations of 300–400 ms, the performance level did not yet differ significantly from the longest stimulus duration (electronic supplementary material, figure S1c) and was still above the 50% threshold, corresponding to two times the chance level (=25%). These results indicate that durations of 300–400 ms should ensure reliable well-above chance level performance,
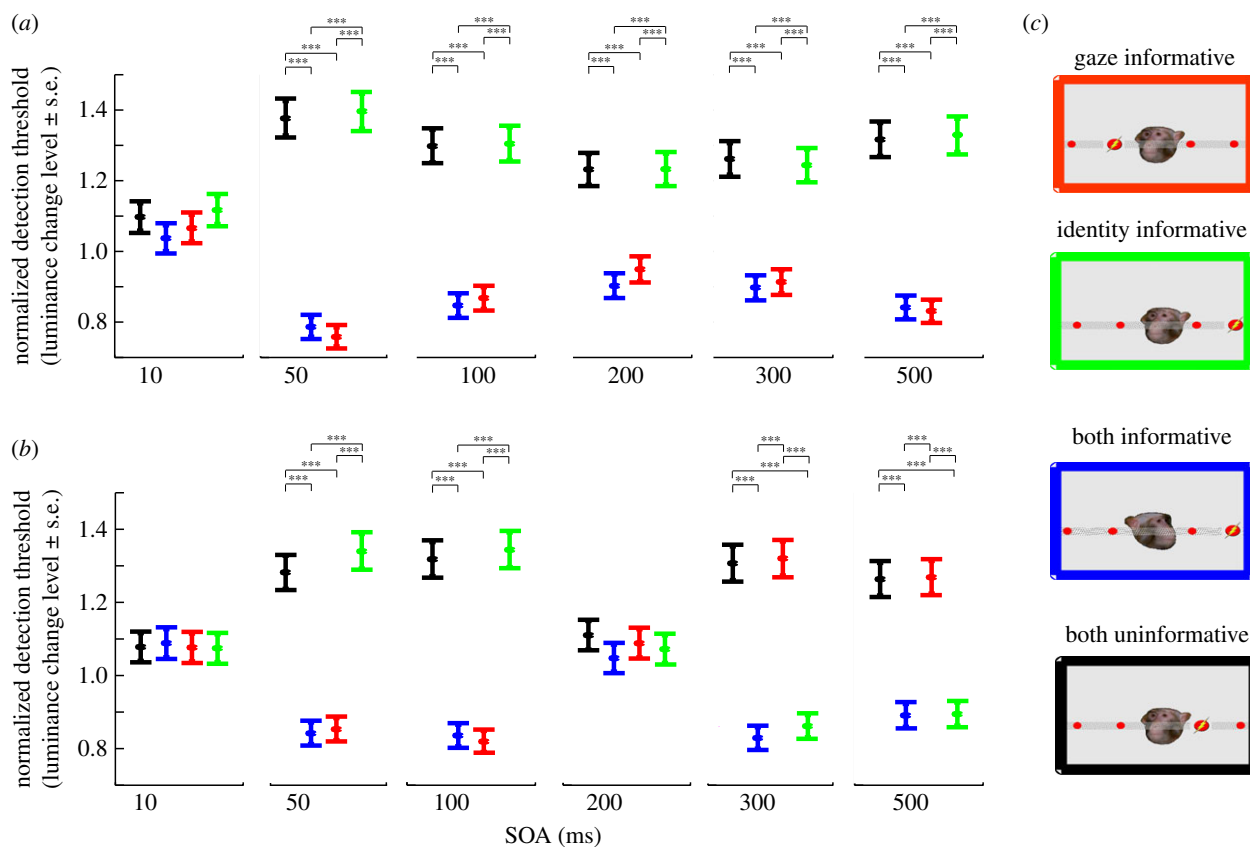
Figure 2. Results of probing the time course of attentional shifts guided by head-gaze or facial identity. Mean normalized detection thresholds as function of SOA for the four congruency categories if the initial rule was head-gaze following ('gaze following context', a) or identity matching ('identity matching context', b). Bars indicate standard errors of the bootstrapping distributions. Three-way ANOVA with the factors context, condition and SOA showed a significant effects of condition ($F_{2.76,2204} = 216$, $p < 0.001$), context × condition interaction ($F_{2.8,2245} = 43$, $p < 0.001$), SOA × condition interaction ($F_{12,9628} = 23$, $p < 0.001$) and context × SOA × condition interaction: $F_{12,9526} = 9.8$, $p < 0.001$. Separate ANOVAs for each context and SOA showed significant effects of condition for each SOA ($p < 0.001$) with the exception of a SOA of 10 ms (gaze $F_{2.4,1908} = 0.981$, $p = 0.387$; identity $p = 0.998$). The results of post hoc comparisons (Bonferroni corrections) are indicated by asterisks: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. No significant differences were found for SOA = 10 ms. (c) Cartoons recapitulating the four congruency categories (see figure 1c for explanation).

but probably only by mobilizing all attentional resources. As in subsequent experiment C, probing the time course of attentional shifts, the stimulus duration had to be as short as possible without jeopardizing performance, 300 was chosen in M1 and M2 for gaze following, M1 for identity matching and 400 ms was chosen in M2 for identity matching as the optimal duration.

In experiment C, we used the subjects' performance in a luminance detection task, embedded in superordinated tasks of gaze following or identity matching ('gaze following context' versus 'identity matching context'), to identify the spatial and temporal location of their attentional shifts. We assumed that the subjects' sensitivity for the luminance change (detection threshold) would be higher, if a preceding gaze or identity cue had drawn their attention to the luminance change location. In the gaze following context, luminance change sensitivities were significantly better for any SOA exceeding 10 ms if the gaze cue was informative, compared to uninformative ones (figure 2a). The identity cue, which according to the prevailing rule was irrelevant, did not influence luminance detection performance. On the other hand, in the facial identity context, gaze direction—now the irrelevant cue—clearly mattered: for SOAs of 50 and 100 ms, luminance change detection was significantly improved if gaze was coincidentally directed at the luminance change target as compared to trials in which this was not the

case, again irrespective of the target to which facial identity pointed. The perceptual influence of the relevant identity cue became apparent only later, at SOAs of 300 and 500 ms: it facilitated performance independent of the gaze cue (figure 2b). To summarize, seen gaze direction prompts an early shift of attention, independent of whether gaze following is called for by the prevailing rule or not. Inappropriate shifts of attention elicited by head-gaze are corrected only later.

Are these early quasi-automatic shifts of attention prompted by head-gaze directed at individual spatial targets? Alternatively, perceived gaze direction might provide an early hemifield advantage, boosting luminance change detection at any target in the hemifield identified by gaze direction while impeding detection in the other one. As shown in figure 3, for both contexts, gaze direction influenced the perceptual threshold only if gaze fell directly onto the target undergoing a change in luminance, suggesting that head-gaze following is indeed spatially precise.

Does the occurrence of an intervening luminance change, expected to attract attention, interfere with the original plan to shift attention, based on the initial cue? In order to obtain an answer, we tested if the performance level was correlated with luminance change level at locations defined either by the gaze or the identity of the portrait. A preceding luminance change had clear effects on the performance in both contexts (electronic supplementary material, figure S2). In support of
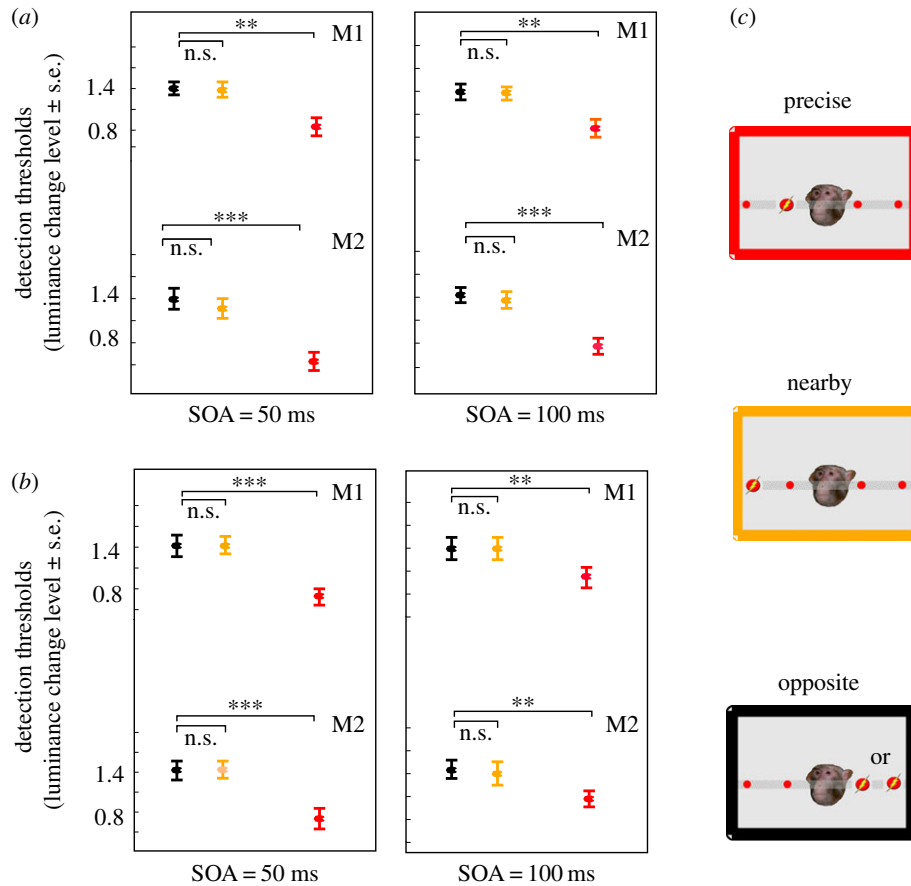
Figure 3. Spatial precision of head-gaze-induced shifts of attention. Mean detection thresholds for luminance changes $\pm$ s.e. for SOAs of 50 and 100 ms for three different trial categories, plotted separately for the two monkeys (M1 = monkey 1, M2 = monkey 2) and the gaze following (*a*) and the identity matching contexts (*b*). Pairwise post hoc statistical comparisons (Bonferroni corrections) are indicated by brackets and the significance level corresponds to the conventions described in figure 2. (One-way ANOVA for each subject and SOA revealed significant effect of precision condition, $p < 0.001$.) (*c*) Cartoons, explaining the three precision categories based on combinations of head-gaze and luminance change location: the 'precise' (red frame), 'nearby' (orange frame) and the 'opposite' (black frame).

the idea of automaticity of gaze following, we found that if the luminance change target and the gaze target positions were spatially congruent, the luminance change had a facilitating effect on the gaze following performance (electronic supplementary material, figure S2*a*, red symbols). Interestingly, it had an impeding effect on the identity matching performance (electronic supplementary material, figure S2*b*, red symbols). These results indicate that luminance changes as well as head-gaze are powerful bottom up attentional cues, involuntarily capturing attention.

Finally, we studied the consequences of having to switch rules for the luminance detection performance. In view of the much stronger compellingness of the head-gaze than the identity cue, we expected that having to switch from head-gaze following to identity matching might come with a larger cost. In order to assess the cost of switching between rules, we asked if luminance change detection thresholds differed between trials in which the given task rule (head-gaze or face matching rule) was the same as in the preceding trial ('repetition trials') or different ('switch trials'). To this end, trials were sorted into separate pools characterized by the presence or absence of a task rule switch and its direction (i.e. from head-gaze following to identity matching or vice versa: four variants), separately for the third shortest and the longest SOA (100 versus 500 ms) and for the type of task that the monkey was asked to carry out: identity matching or gaze following. The short SOA was chosen as it prompted automatic, rule-independent shifts of attention

guided by head-gaze. On the other hand, the longest SOA had shown consistent shifts of attention based on the head-gaze as well as on the identity rule, in the latter case no longer influenced by head-gaze. Surprisingly, the statistical analysis (see the electronic supplementary material, figure 3, for details) failed to reveal significant switch cost effects for the 100 ms SOA. Also for the 500 ms SOA, we did not observe a switch cost effect in the sense that the need to switch the rule would have led to poorer performance. Actually, we found a significantly better performance for switch trials compared to repetition trials selectively for the combination of the identity matching rule to be applied and the gaze cue—to be ignored—being informative: in this constellation, the monkey is asked to use identity information although gaze determines the site of the luminance change. The fact that the threshold is better for switch trials reflects the fact that the monkey has not yet managed to fully suppress the gaze following rule valid in the preceding trial, a deficiency which we would understand as an interference effect. This interference effect is not seen for the 100 ms SOA, simply because for this short SOA there is not yet a significant shift of attention based on identity that could be disrupted by gaze.

4. Discussion

We compared the ability of two types of social cues to shift monkey's spatial attention. The first cue was peer head-gaze

direction, the second one peer facial identity. Head-gaze direction identifies a distinct spatial location as an intersection of the head-gaze direction vector and a potential object of interest. It should provide salient geometric information that in principle is accessible without prior learning. On the other hand, the assignment of particular facial identities to distinct spatial positions is quite unnatural and requires the learning of arbitrary associations. Our monkeys learned to cope with both tasks as indicated by comparably low error rates. Moreover, they learned a third task, requiring a change of plan from shifting attention to the target identified by the social cue at stake, to detecting a luminance change of any of the four targets in the set, not necessarily the one singled out by the social cue. In other words, this task variant required a switch from spatial to object-based attention, allowing us to study the interactions of the two. Reliable luminance change detection thresholds could be determined in most, albeit not all sessions (see Material and methods). The fact that the logistic fits used to pinpoint thresholds failed to be significant in a minority of sessions may have been a reflection of insufficient motivation to cope with the complex task in these sessions. Our results show that both social cues reallocated spatial attention, which improved luminance change detection, if the focus of spatial attention coincidentally overlapped with the luminance change target. In case of facial identity, the shift of spatial attention required at least about 300 ms as indicated by optimal SOAs of 300–500 ms. On the other hand, the shift of attention prompted by the gaze cue was much faster as SOAs as short as 50 ms improved luminance detection. The fact that a clear perceptual benefit from the gaze cue was visible also for the longest SOA tested indicates that once allocated the attentional spotlight stayed out. Finally, our data clearly indicate that only the shift of attention guided by facial identity can be fully suppressed if the other social cue is known to be the relevant one. On the other hand, the gaze signal prompted the early shift of attention to the gaze contingent target, regardless of its relevance. If the task rule demanded relying on facial identity, this early misallocation of spatial attention started to become corrected only later, at about 300 ms. These findings clearly suggest that head-gaze following of rhesus monkeys is a fast and automatic, quasi-reflex-like behaviour. Moreover, our data show that a head-gaze cue works as distractor, interfering with the execution of a different task, called for by the prevailing rule. Previous studies have shown that in general monkeys' ability to focus on the task at hand is comparatively poor: their performance is significantly affected by irrelevant stimulus features (high interference costs, see [22] for review). The interference of the head-gaze cue (and the head-gaze rule) on a subsequent face matching task is in line with these previous studies and in accordance with the notion of compellingness, little controllabilty and automaticity of head-gaze following in monkeys. The absence of relevant switch cost effects in our experiments is in accordance with previous studies which have argued that monkeys may in general be less prone to switch costs than humans (see [22] for review). Automaticity of monkeys' gaze following was also suggested by a study by Deaner & Platt [12] in which monkeys saw a conspecific's head oriented to the right or left. Without any specific instruction or behavioural incentive, the observing monkeys tended to shift their attention into the hemifield pointed to by the portraited monkeys. Our results clearly indicate that this head-gaze following reaction is very fast and, moreover, that it cannot be fully suppressed, even if this was demanded

by the prevailing rule and, most importantly, that it is indeed geometric. In other words, monkeys are able to use the vector of others' head-gaze direction to pinpoint a specific spatial position, given that it meets an object of potential interest. On the other hand, the ability of our monkeys to use facial identity to deploy spatial attention seems to rely on a general purpose association machinery, as suggested by both its sluggishness and its perfect controllability. Also humans may exploit features of centrally presented non-social cues in order to generate shifts of attention into the periphery. An example are non-directional symbols that acquire their spatial connotation based on learning such as directional words [23], numbers [24] or even word categories paired arbitrarily with target locations [25]. These verbal cues cause shifts of spatial orientation which are rather slow and which can be suppressed if needed [26]. This is in full accordance with our findings on attentional shift of rhesus monkeys guided by facial identity. On the other hand, centrally presented symbols with immediate spatial value like pointing fingers [27], a pointing tongue [28] or arrows prompt attentional shifts in humans at short latencies, comparable to those of eye-gaze cues [29]. This similarity between shifts of spatial attention evoked by eye-gaze and by other cues providing spatial information in humans might suggest that both are based on the spatial compatibility between the cue and the target. Yet, humans have much more control over their focus of attention when being exposed to arrows than to eye-gaze cues [30]. Therefore, what makes human eye-gaze following special is its resistance to top-down influences. As shown in this study, a similar resistance to top-down control characterizes monkeys' head-gaze following. However, this is not to say that gaze following in monkeys and man is an obligatory reflex lacking any top-down modulation. In fact, we observed suppression, yet only after an initial transient shift of attention prompted by the gaze cue. Actually, work on human eye-gaze following suggests that it may consist of an early, hardly controllable component and a later component that is subject to top-down control, ensuring that it is released only if useful [31]. Our results on monkeys' head-gaze following favour a comparable sequence of an uncontrollable early shift of attention that takes place no matter if demanded by the task or not and a later instruction contingent voluntary shift component. The balance between these components might vary with factors like facial expression [32], social status, social preferences [33–35] or mental states attributed to the sender [36].

Finally, we emphasize yet another close correspondence between human eye-gaze following and monkeys' head-gaze following, namely the fact that both allocate attention to distinct locations in the visual field, rather than redistributing resources between the two hemifields. Actually, as shown by several studies, human eye-gaze following is able to single out particular locations with a precision of a few degrees only [37,38] and to confine the attentional spotlight to an area as small as 2–4° [39,40]. The paradigm used in our study of head-gaze following in monkeys did not allow us to precisely gauge the boundaries of the attentional spotlight. Yet, it allowed us to ascertain that it singles out restricted areas in a given hemifield, probably not larger than 5°. This result supports the notion that monkeys' head-gaze following is geometric, much like human eye-gaze following.

In summary, monkeys' head-gaze shares key features of human eye-gaze following, namely, automaticity and swiftness, limited executive control and the consideration of the

geometry of the triadic constellation defined by the positions of the observer, the demonstrator's gaze and his/her potential object of interest. This close correspondence adds to converging evidence for a conserved neural circuitry with very similar properties in humans and monkeys. This conclusion is actually very much in line with results from recent fMRI, which have implicated a distinct patch of cortex in the superior temporal sulcus activated by eye-gaze following in humans [11] and head-gaze following in monkeys [19], not only sharing a common general topography but also comparable relationships to the face patch system (monkeys [19] and human [41]).

References

1. Emery NJ. 2000 The eyes have it: the neuroethology, function and evolution of social gaze. Neurosci. Biobehav. Rev. 24, 581–604. (doi:10.1016/S0149-7634(00)00025-7)

2. Driver J, Davis G, Ricciardelli P, Kidd P, Maxwell E, Baron-Cohen S. 1999 Gaze perception triggers reflexive visuospatial orienting. Vis. Cogn. 6, 509–540. (doi:10.1080/135062899394920)

3. Friesen CK, Kingstone A. 1998 The eyes have it! reflexive orienting is triggered by nonpredictive gaze. Psychon. Bull. Rev. 5, 490–495. (doi:10.3758/BF03208827)

4. Posner MI, Cohen Y. 1984 Components of visual orienting. Atten. Perform. X 32, 531–556.

5. Posner MI, Rafal RD, Choate L, Vaughan J. 1985 Inhibition of return: neural mechanisms and function. Cogn. Neuropsychol. 2, 211–228. (doi:10.1080/02643298508252866)

6. Frischen A, Tipper SP. 2004 Orienting attention via observed gaze shift evokes longer term inhibitory effects: implications for social interactions, attention, and memory. J. Exp. Psychol. Gen. 133, 516–533. (doi:10.1037/0096-3445.133.4.516)

7. Frischen A, Smilek D, Eastwood JD, Tipper SP. 2007 Inhibition of return in response to gaze cues: evaluating the roles of time course and fixation cue. Vis. Cogn. 15, 881–895. (doi:10.1080/13506280601112493)

8. Batki A, Baron-Cohen S, Wheelwright S, Connellan J, Ahluwalia J. 2000 Is there an innate gaze module? Evidence from human neonates. Infant. Behav. Dev. 23, 223–229. (doi:10.1016/S0163-6383(01)00037-6)

9. Hood BM, Willen JD, Driver J. 1998 Adult's eyes trigger shifts of visual attention in human infants. Psychol. Sci. 9, 131–134. (doi:10.1111/1467-9280.00024)

10. Akiyama T, Kato M, Muramatsu T, Saito F, Umeda S, Kashima H. 2006 Gaze but not arrows: a dissociative impairment after right superior temporal gyrus damage. Neuropsychologia 44, 1804–1810. (doi:10.1016/j.neuropsychologia.2006.03.007)

11. Materna S, Dicke PW, Thier P. 2008 Dissociable roles of the superior temporal sulcus and the intraparietal sulcus in joint attention: a functional magnetic resonance imaging study. J. Cogn. Neurosci. 20, 108–119. (doi:10.1162/jocn.2008.20008)

12. Deaner RO, Platt ML. 2003 Reflexive social attention in monkeys and humans. Curr. Biol. 13, 1609–1613. (doi:10.1016/j.cub.2003.08.025)

13. Lorincz EN, Baker CI, Perrett DI. 1999 Visual cues for attention following in rhesus monkeys. Curr. Psychol. Cogn. 18, 973–1003.

14. Ferrari PF, Kohler E, Fogassi L, Gallese V. 2000 The ability to follow eye gaze and its emergence during development in macaque monkeys. Proc. Natl Acad. Sci. USA 97, 13 997–14 002. (doi:10.1073/pnas.250241197)

15. Kobayashi H, Kohshima S. 1997 Unique morphology of the human eye. Nature 387, 767–768. (doi:10.1038/42842)

16. 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. J. Hum. Evol. 52, 314–320. (doi:10.1016/j.jhevol.2006.10.001)

17. Marciniak K, Atabaki A, Dicke PW, Thier P. 2014 Disparate substrates for head gaze following and face perception in the monkey superior temporal sulcus. Elife 3, e03222. (doi:10.7554/eLife.03222)

18. Lieberman HR, Pentland AP. 1982 Microcomputer-based estimation of psychophysical thresholds: the best PEST. Behav. Res. Meth. Inst. 14, 21–25. (doi:10.3758/BF03202110)

19. Kingdom FAA, Prins N. 2010 Psychophysics: a practical introduction. London, UK: Academic Press.

20. McKee SP, Klein SA, Teller DY. 1985 Statistical properties of forced choice psychometric functions: implications of probit analysis. Percept. Psychophys. 37, 286–298. (doi:10.3758/BF03211350)

21. Foster D, Bischof WF. 1991 Thresholds from psychometric functions: superiority of bootstrap to incremental and probit variance estimators. Psychol. Bull. 109, 152–159. (doi:10.1037/0033-2909.109.1.152)

22. Stoet G, Snyder LH. 2003 Executive control and task-switching in monkeys. Neuropsychologia 41, 1357–1364. (doi:10.1016/S0028-3932(03)00048-4)

23. Hommel B, Pratt J, Colzato L, Godijn R. 2001 Symbolic control of visual attention. Psychol. Sci. 12, 360–365. (doi:10.1111/1467-9280.00367)

24. Fischer MH. 2003 Spatial representations in number processing: evidence from a pointing task. Vis. Cogn. 10, 493–508. (doi:10.1080/13506280244000186)

25. Lambert AJ, Sumich AL. 1996 Spatial orienting controlled without awareness: a semantically based implicit learning effect. Q. J. Exp. Psychol. A 49, 490–518. (doi:10.1080/713755621)

26. Galfano G, Rusconi E, Umiltà C. 2006 Number magnitude orients attention, but not against one's will. Psychol. Bull. Rev. 13, 869–874. (doi:10.3758/BF03194011)

27. Langton SRH, Bruce V. 2000 You must see the point: automatic processing of cues to the direction of social attention. J. Exp. Psychol. Hum. Percept. Perform. 26, 747–757. (doi:10.1037/0096-1523.26.2.747)

28. Downing P, Dodds CM, Bray D. 2004 Why does the gaze of others direct visual attention? Vis. Cogn. 11, 71–79. (doi:10.1080/13506280344000220)

29. Tipples J. 2002 Eye gaze is not unique: automatic orienting in response to uninformative arrows. Psychon. Bull. Rev. 9, 314–318. (doi:10.3758/BF03196287)

30. Friesen CK, Ristic J, Kingstone A. 2004 Attentional effects of counterpredictive gaze and arrow cues. J. Exp. Psychol. Hum. Percept. Perform. 30, 319–329. (doi:10.1037/0096-1523.30.2.319)

31. Hill JL, Patel S, Gu X, Seyedali NS, Bachevalier J, Sereno AB. 2010 Social orienting: reflexive versus voluntary control. Vis. Res. 50, 2080–2092. (doi:10.1016/j.visres.2010.07.020)

32. Teufel C, Gutmann A, Pirow R, Fischer J. 2010 Facial expressions modulate the ontogenetic trajectory of gaze-following among monkeys. Dev. Sci. 13, 913–922. (doi:10.1111/j.1467-7687.2010.00956.x)

8

rspb.royalsocietypublishing.org   Proc. R. Soc. B 282: 20151020

9

rspb.royalsocietypublishing.org   *Proc. R. Soc. B* **282**: 20151020

33. Shepherd SV, Deaner RO, Platt ML. 2006 Social status gates social attention in monkeys. *Curr. Biol.* **16**, R119–R120. (doi:10.1016/j.cub.2006.02.013)

34. Dodd MD, Hibbing JR, Smith KB. 2011 The politics of attention: gaze-cuing effects are moderated by political temperament. *Atten. Percept. Psychophys.* **73**, 24–29. (doi:10.3758/s13414-010-0001-x)

35. Liuzza MT, Cazzato V, Vecchione M, Crostella F, Caprara GV, Aglioti SM. 2011 Follow my eyes: the gaze of politicians reflexively captures the gaze of ingroup voters. *PLoS ONE* **6**, e25117. (doi:10.1371/journal.pone.0025117)

36. Teufel C, Fletcher PC, Davis G. 2010 Seeing other minds: attributed mental states influence perception. *Trends Cogn. Sci.* **14**, 376–382. (doi:10.1016/j.tics.2010.05.005)

37. Bock S, Dicke PW, Thier P. 2008 How precise is gaze following in humans? *Vis. Res.* **48**, 946–957. (doi:10.1016/j.visres.2008.01.011)

38. Symons LA, Lee K, Cedrone CC, Nishimura M. 2004 What are you looking at? Acuity for triadic eye gaze. *J. Gen. Psychol.* **131**, 451–469.

39. Vuilleumier P. 2002 Perceived gaze direction in faces and spatial attention: a study in patients with parietal damage and unilateral neglect. *Neuropsychologia* **40**, 1013–1026. (doi:10.1016/S0028-3932(01)00153-1)

40. Wiese E, Zwickel J, Müller HJ. 2013 The importance of social cotext information for the spatial specificity of gaze cueing. *Atten. Percept. Psychophys.* **75**, 967–982. (doi:10.3758/s13414-013-0444-y)

41. Marquardt K, Ramezanpour H, Dicke PW, Thier P. 2015 *Following the eye gaze of others activates a patch of cortex that is not part of the 'face patch' system. Neuroscience 2015 abstracts*. Chicago, IL: Society for Neuroscience.
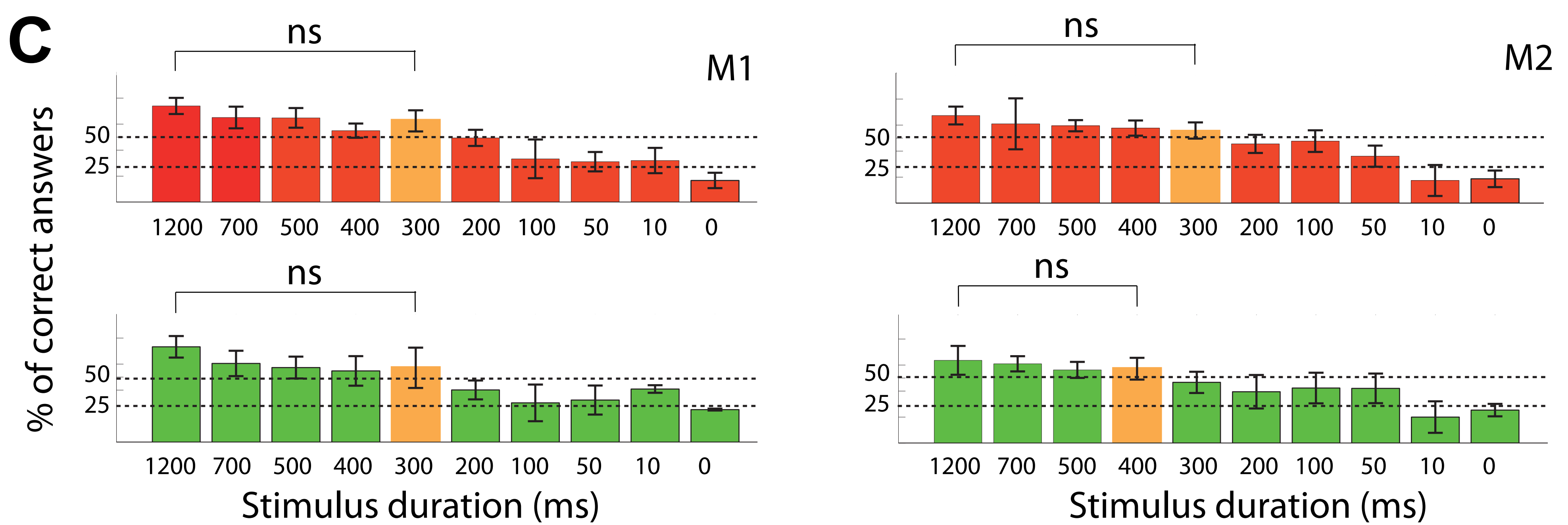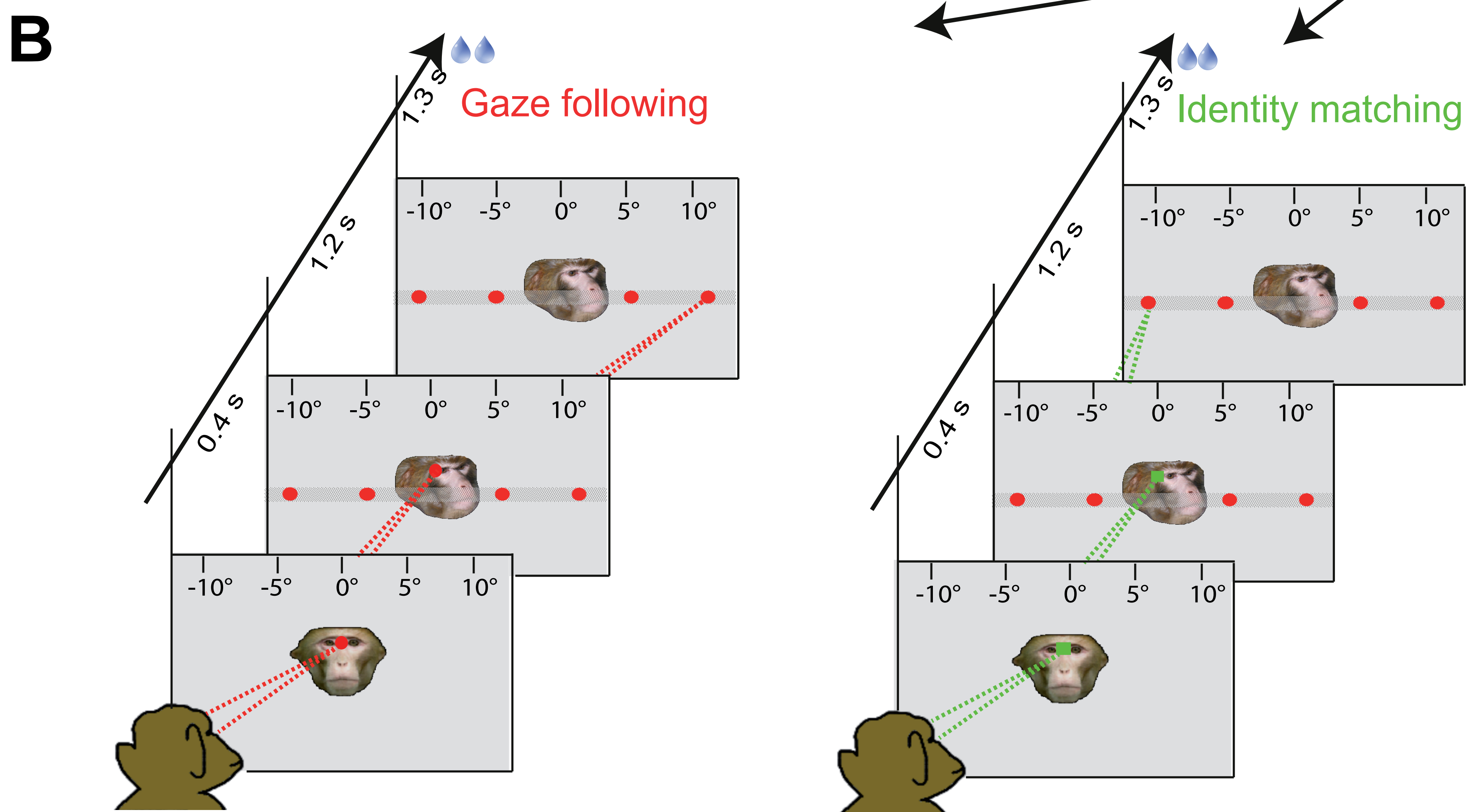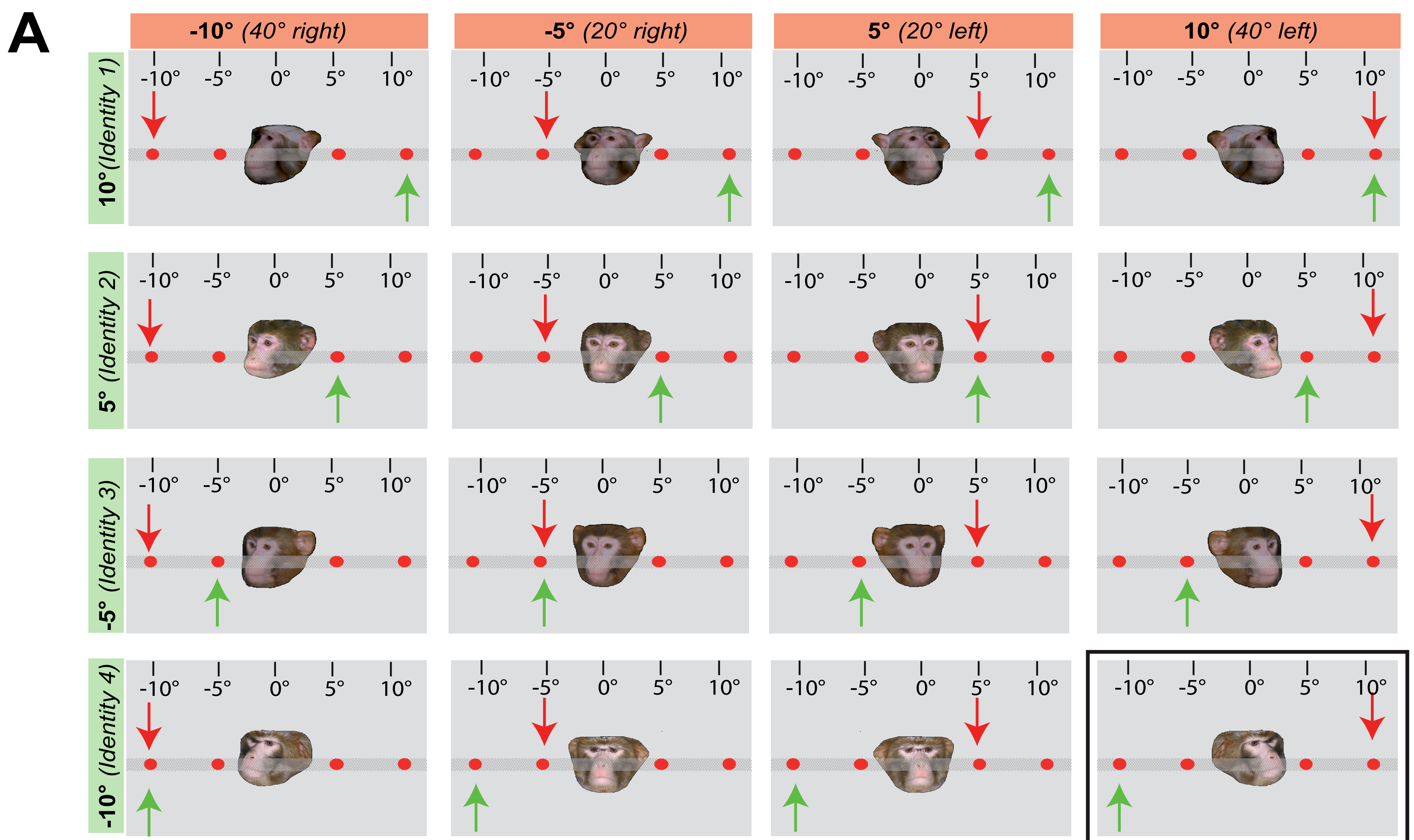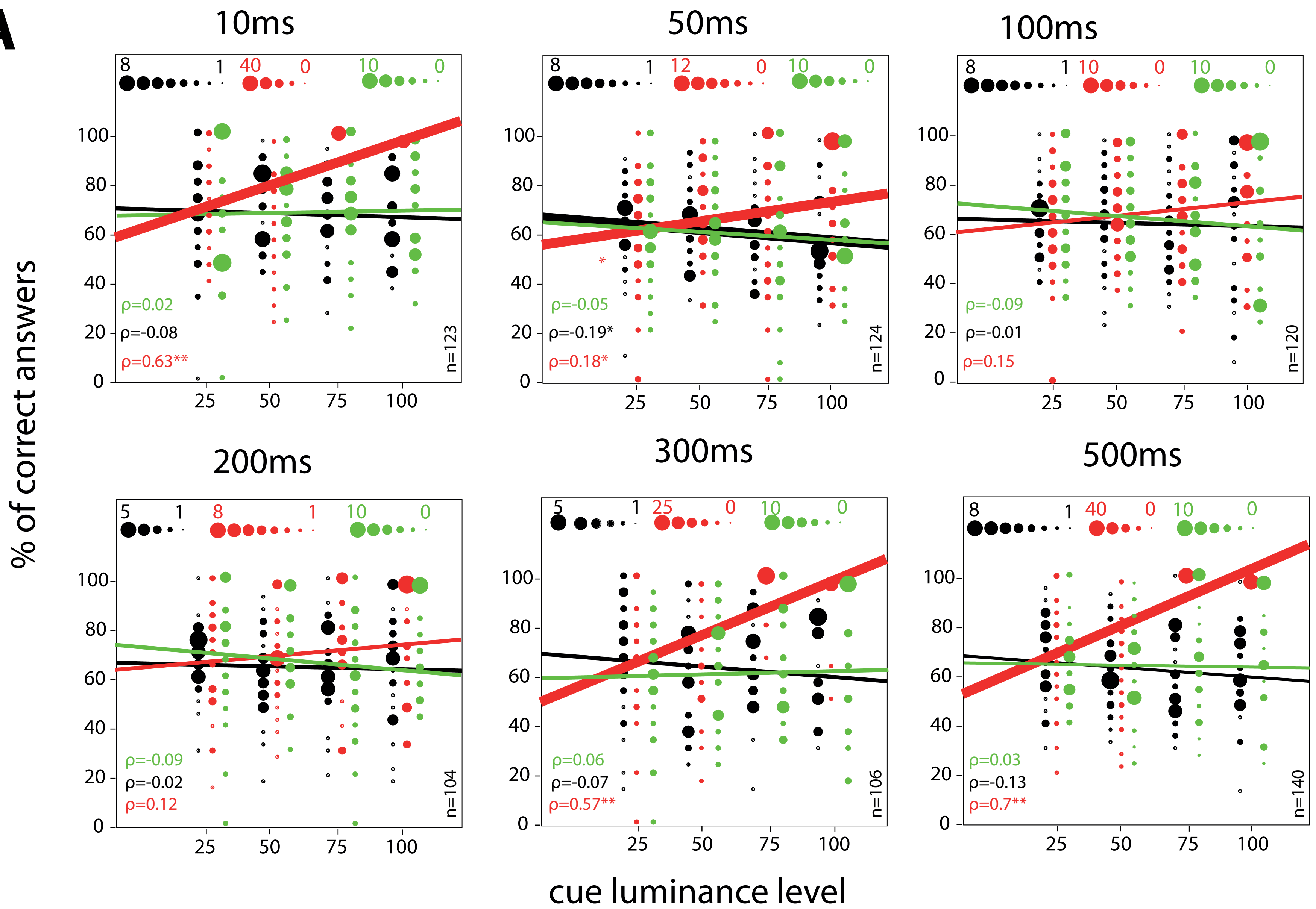
Gaze following and identity matching paradigms

A Stimuli. The same 16 portraits showing 4 individual monkeys from 4 different views used in the gaze following and identity-matching tasks arranged by identity (columns) and head orientation (rows). The arrows point to the correct gaze following (red) or identity matching (green) target. Arrows and the scale were not visible during the experiment. Portraits and target bar were presented on an otherwise black background (here shown as gray for better visualization).
B. Sequence of events. Exemplary gaze following (left) and identity matching (right) trials. **C**. Results of the stimulus duration adjustment experiment. Mean performance level (bars indicate standard errors) as a function of portrait duration for gaze following (red) and identity matching (green). The stimulus duration chosen for the subsequent luminance detection experiments, ensuring a performance level 50% is shown in orange. M1=monkey 1, M2=monkey 2, ns= not significant. One-way ANOVA with the factor stimulus duration, done separately for each monkey and task revealed significant effect of stimulus duration (M2identity: $F(9,50)=13.8$, $p<0.001$; M2gaze: $F(9,40)=17.2$, $p<0.001$; M1 identity: $F(9,40)=12.4$, $p<0.001$; M1 gaze: $F(9,40)=25.6$, $p<0.001$; ns: no significant differences (M1gaze: $p=1$, M1id: $p=0.635$, M2gaze: $p=1$, M2id: $p=1$) in posthoc tests after Bonferroni corrections.

Supplemental Figure 2
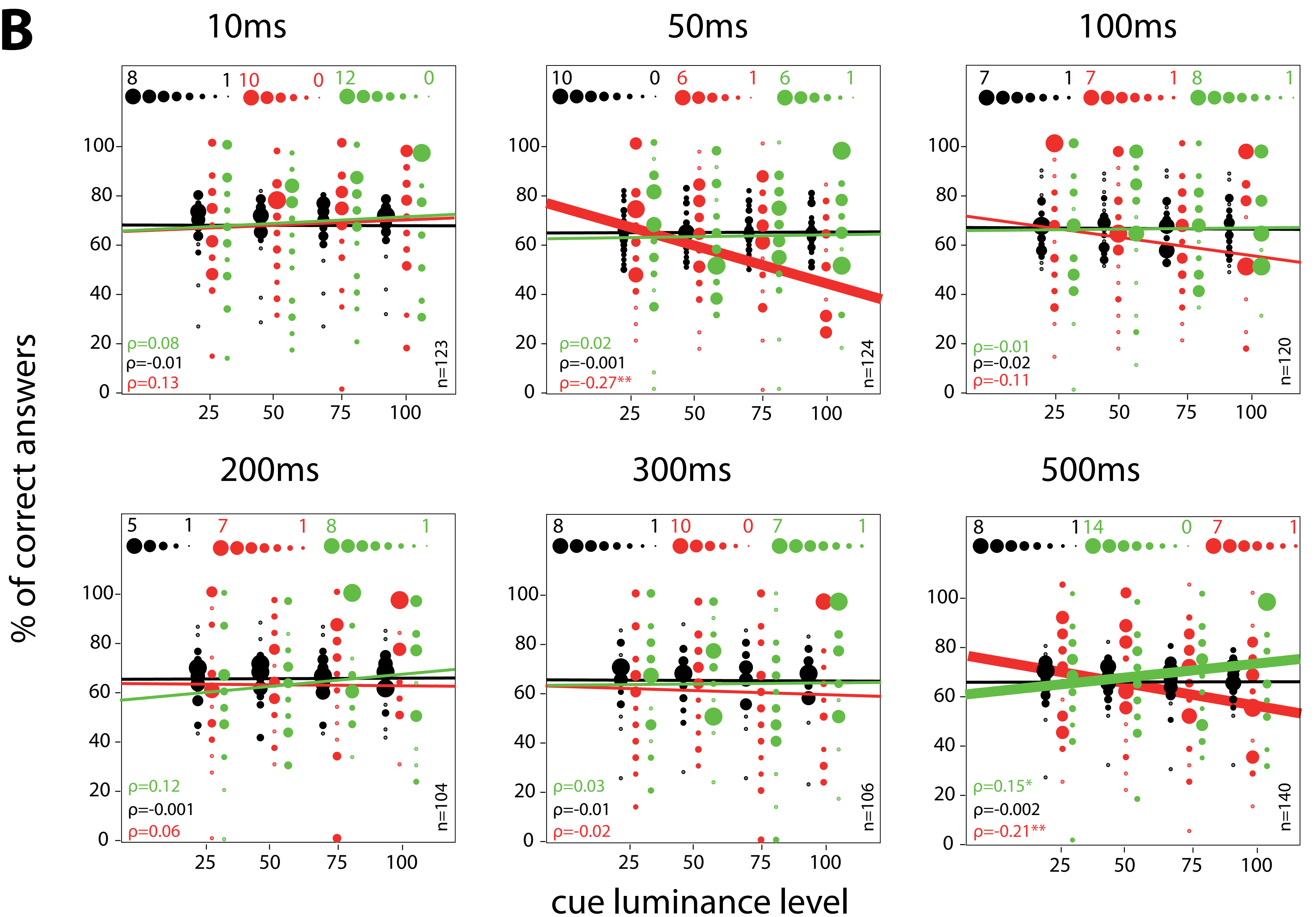
The effect of an intervening luminance change on head-gaze following (A) and identity-matching (B). Performance level as a function of the luminance change level is drawn as dots separately for the congruency categories: gaze informative (red), identity informative (green), and neither of the two informative (black) as well as the various SOAs .The number in the lower right corner (n) of each plot specifies the overall number of experimental sessions. The diameter of the individual dot reflects the number of observations per case. The lines represent linear regressions. The correlation coefficients (Spearman's rho, ρ) are depicted in the lower left corner for the respective class together with symbols reflecting the significance levels (p<0.001: ***; p<0.01: **, p<0.05: *). Significant regressions are emphasized by bold lines.
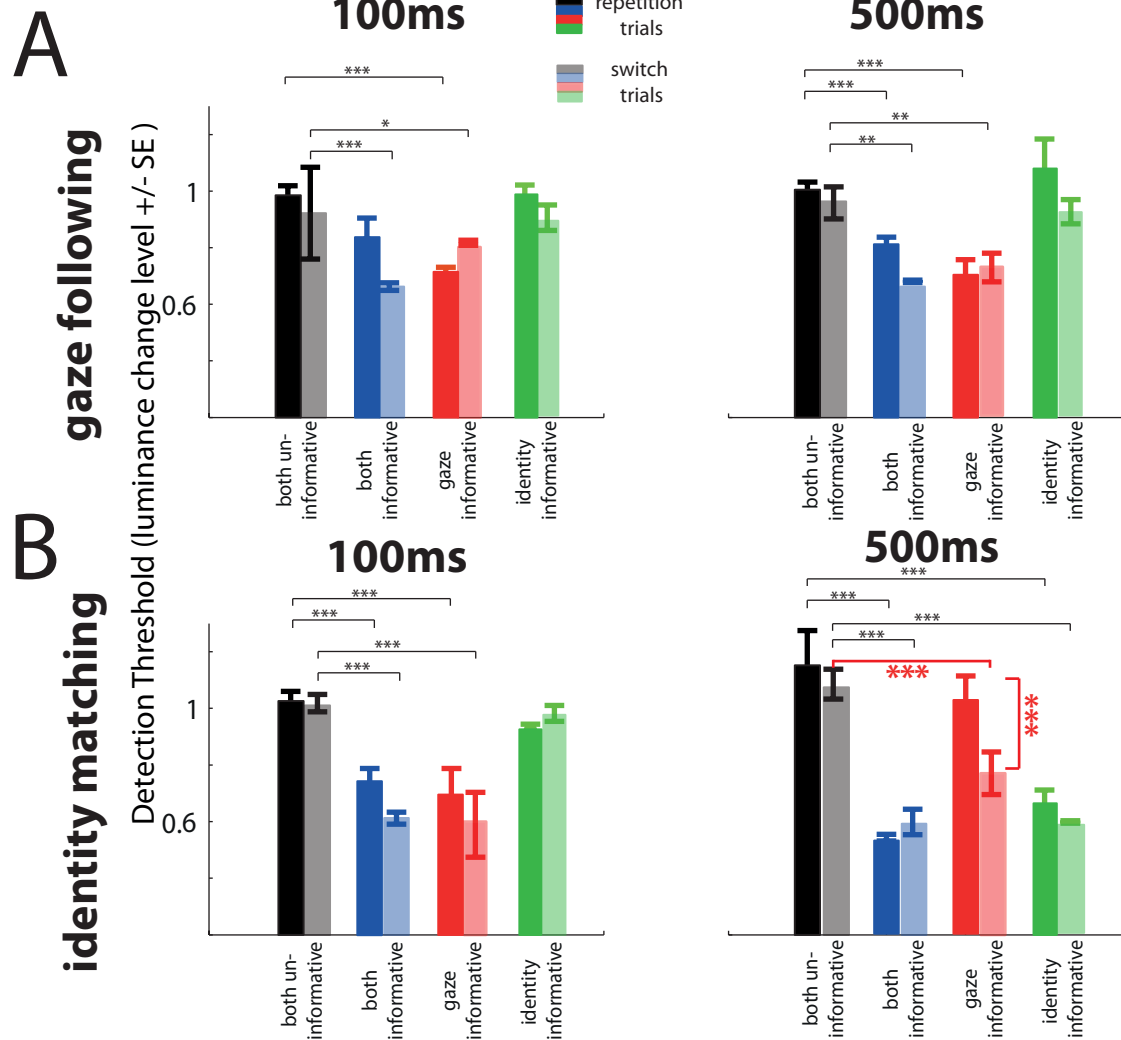
Switching costs in luminance detection. Luminance detection thresholds for 4 conditions, distinguished by the congruency of spatial information provided by head gaze, facial identity with the actual target location (black, blue, red and green indicate the various constellations as described in Fig. 2, calculated for trials in which the gaze following (A) and for trials in which the identity matching rule (B) governed and separately for an SOA of 100ms (left) and 500ms (right). Bars plotted in light colours indicate 'switch' trials (the task rule in the preceding trial was opposite to the rule in a current one), and those plotted in dark colours 'repetition' trials (the task rule of the previous trial is repeated in a current trial). A 4 way repeated measures ANOVA with the factors SOA (100ms, 500ms), task rule (head-gaze vs. identity matching), trial type (repetition vs. switch trials) and congruency condition (both cues uninformative, both informative, gaze informative, identity informative) revealed a significant interaction of trial type x SOA x congruency condition (F[6,16.5]=3.03, p=0.006). A 3 way repeated measures analysis, sparing the SOA factor, applied separately to each oft he two individual SOA conditions showed a significant trial type x task rule x congruency condition interaction for the 500ms SOA only (F[2.3,14.3]=5.3, p=0.004). Significant pairwise comparisons (with Bonferroni corrections) are indicated by black asterisks (*->p<0.05, **->p<0.01, ***->p<0.001). Comparisons shown in black are consistent with the effect pattern summarized in Fig.2, those in red denote that significant head-gaze cue improvements were found only in switch trials but not in repetition trials.

Supplemental Table 1

Condition of psychometric function		M1					M2				
		detection threshold	slope	pDev	Dev	std	detection threshold	slope	pDev	Dev	std
SOA=10ms	incongruent	0.99	12.6	0.586	1.018	0.537	1.031	14.751	0.381	5.232	1.547
	both congruent	0.894	0.919	0.809	0.481	1.197	1.051	1.147	0.748	0.684	1.840
	gaze-only congruent	1	0.005	0.330	26.820	0.009	1.069	6.412	0.848	4.410	0.836
	identity-only congruent	1.004	4.250	0.793	0.069	0.307	1.127	3.882	0.856	0.297	1.171
SOA=50ms	incongruent	1.152	1.873	0.826	0.421	1.013	1.094	3.987	0.736	0.648	1.705
	both congruent	0.753	1.712	0.831	0.434	0.919	0.744	3.121	0.806	1.035	0.153
	gaze-only congruent	0.684	2.125	0.918	0.216	0.226	0.825	5.849	0.951	5.849	0.079
	identity-only congruent	1.106	1.702	0.938	0.145	1.357	1.165	3.432	0.710	1.465	1.503
SOA=100ms	incongruent	1.119	3.672	0.516	22.461	0.889	1.157	2.415	0.470	0.560	1.299
	both congruent	0.778	2.435	0.849	24.739	1.038	0.820	2.114	0.599	1.123	1.792
	gaze-only congruent	0.856	1.658	0.375	29.869	0.814	0.866	4.699	0.954	0.317	0.142
	identity-only congruent	1.081	2.752	0.866	0.333	1.494	1.120	2.575	0.393	1.878	1.260
SOA=200ms	incongruent	1.119	3.435	0.870	0.290	0.658	0.976	6.585	0.831	0.043	1.689
	both congruent	0.921	2.731	0.619	1.749	0.407	0.818	6.872	0.884	0.274	0.764
	gaze-only congruent	0.931	1.654	0.936	0.135	0.661	0.879	5.698	0.269	1.267	0.083
	identity-only congruent	1.125	1.052	0.980	0.039	1.278	1.009	6.320	0.642	0.219	0.748
SOA=300ms	incongruent	1.143	2.303	0.333	1.040	0.915	1.087	2.785	0.562	1.352	1.192
	both congruent	0.855	2.977	0.445	1.635	0.437	0.882	3.581	0.573	1.765	0.227
	gaze-only congruent	0.870	2.032	0.945	0.144	0.425	0.870	5.690	0.593	5.248	0.180
	identity-only congruent	1.232	0.982	0.955	0.106	1.694	1.050	4.180	0.786	0.483	0.384
SOA=500ms	incongruent	1.035	1.990	0.730	0.718	0.707	1.073	3.487	0.994	0.021	1.553
	both congruent	0.742	4.496	0.912	0.598	0.079	0.747	1.546	0.832	1.094	0.712
	gaze-only congruent	0.719	2.402	0.975	0.215	0.257	0.781	10.356	0.575	1.271	0.048
	identity-only congruent	0.976	1.769	0.33	2.277	1.615	1.066	1.584	0.702	0.698	1.613

Supplemental Table 2

		M1					M2				
Condition of psychometric function		detection threshold	slope	pDev	Dev	std	detection threshold	slope	pDev	Dev	std
SOA=10ms	incongruent	1.013	49.211	0.858	20.289	0.126	1.192	4.614	0.348	3.845	2.693
	both congruent	1.031	1.401	0.638	0.913	0.705	1.126	2.127	0.192	11.273	2.914
	gaze-only congruent	0.989	9.401	0.930	0.026	0.205	1.171	2.728	0.067	4.892	2.053
	identity-only congruent	0.993	2.527	0.938	0.009	0.384	1.232	1.555	0.779	0.085	2.626
SOA=50ms	incongruent	1.151	1.134	0.671	8.187	1.882	0.973	24.049	0.517	1.455	0.032
	both congruent	0.759	1.927	0.606	0.297	0.399	0.775	5.198	0.291	2.388	0.100
	gaze-only congruent	0.731	2.353	0.990	0.020	0.406	0.749	2.139	0.119	39.817	1.406
	identity-only congruent	1.204	1.049	0.718	0.174	1.810	1.051	2.565	0.659	0.208	0.938
SOA=100ms	incongruent	1.176	4.409	0.781	1.063	1.011	1.033	4.324	0.795	0.480	1.357
	both congruent	0.794	2.891	0.907	0.613	0.487	0.829	5.080	0.953	0.331	0.102
	gaze-only congruent	0.829	3.949	0.512	1.353	0.139	0.743	2.088	0.905	0.576	1.594
	identity-only congruent	1.235	2.255	0.932	0.175	1.280	1.005	2.377	0.061	5.622	2.154
SOA=200ms	incongruent	1.181	4.681	0.604	0.999	0.986	0.955	4.877	0.397	2.226	0.252
	both congruent	1.069	2.402	0.901	0.222	0.895	0.956	1.742	0.990	0.021	0.374
	gaze-only congruent	0.123	2.264	0.844	0.315	0.680	0.948	3.882	0.242	2.897	0.311
	identity-only congruent	1.063	7.574	0.533	1.238	1.255	0.997	8.370	0.925	0.195	0.129
SOA=300ms	incongruent	1.189	2.831	0.915	0.205	1.254	1.174	4.900	0.776	0.086	1.215
	both congruent	0.780	3.470	0.574	2.146	0.278	1.056	4.902	0.943	0.006	1.245
	gaze-only congruent	1.256	0.938	0.971	0.759	1.905	1.251	3.215	0.601	1.082	1.272
	identity-only congruent	0.940	1.778	0.810	0.485	2.947	1.047	5.356	0.230	3.037	1.149
SOA=500ms	incongruent	0.792	0.860	0.845	0.333	1.206	1.124	2.132	0.991	0.029	0.916
	both congruent	0.411	5.112	0.565	2.368	0.105	0.784	2.701	0.665	1.576	1.499
	gaze-only congruent	0.730	4.481	0.838	0.323	0.106	1.058	2.718	0.806	0.084	1.724
	identity-only congruent	0.495	3.756	0.712	0.726	0.116	0.791	3.657	0.417	2.756	0.119