Effect of distracting faces on visual selective attention in the monkey

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1073/pnas.1420167111">http://dx.doi.org/10.1073/pnas.1420167111</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>National Academy of Sciences (U.S.)</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Tue Dec 11 02:39:02 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/97428">http://hdl.handle.net/1721.1/97428</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Effect of distracting faces on visual selective attention in the monkey

Rogier Landman, Jitendra Sharma, Mriganka Sur, and Robert Desimone

*McGovern Institute for Brain Research and #Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA 02139; #Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA 02129; Department of Radiology, Harvard Medical School, Boston, MA 02115; and *Simons Center for the Social Brain, Cambridge, MA 02139

Contributed by Robert Desimone, November 3, 2014 (sent for review April 20, 2014; reviewed by Luiz Pessoa and Wim Van Duffel)

In primates, visual stimuli with social and emotional content tend to attract attention. Attention might be captured through rapid, automatic, subcortical processing or guided by slower, more voluntary cortical processing. Here we examined whether irrelevant faces with varied emotional expressions interfere with a covert attention task in macaque monkeys. In the task, the monkeys monitored a target grating in the periphery for a subtle color change while ignoring distracters that included faces appearing elsewhere on the screen. The onset time of distracter faces before the target change, as well as their spatial proximity to the target, was varied from trial to trial. The presence of faces, especially faces with emotional expressions interfered with the task, indicating a competition for attentional resources between the task and the face stimuli. However, this interference was significant only when faces were presented for greater than 200 ms. Emotional faces also affected saccade velocity and reduced pupillary reflex. Our results indicate that the attraction of attention by emotional faces in the monkey takes a considerable amount of processing time, possibly involving cortical-subcortical interactions. Intranasal application of the hormone oxytocin ameliorated the interfering effects of faces. Together these results provide evidence for slow modulation of attention by emotional distracters, which likely involves oxytocinergic brain circuits.

Significance

Primates express a natural interest in faces. The viewing of faces with an emotional expression affects emotion circuits in the brain, even when they are not directly attended. This has led to a debate about whether faces attract attention automatically. We tested the influence of emotional faces as irrelevant distracters in an attention task in monkeys. Task performance was most affected when facial expression was threatening, especially when presented for durations longer than 200 ms. We conclude that, in monkeys, emotional distracters do attract attention away from other tasks but not instantly. Administration of the hormone oxytocin reduced the effect. Among the brain systems likely involved are areas where oxytocin receptors are abundant.

Author contributions: R.L., J.S., M.S., and R.D. designed research; R.L. and J.S. performed research; R.L. and J.S. analyzed data; and R.L., J.S., M.S., and R.D. wrote the paper.

Reviewers: L.P., University of Maryland; and W.V.D., Harvard/Massachusetts General Hospital.

The authors declare no conflict of interest.

Author contributions: R.L. and J.S. contributed equally to this work.

6To whom correspondence may be addressed. Email: landman@mit.edu or desimone@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1420167111/-/DCSupplemental.
Results
The monkeys performed a task in which two colored gratings, a target and a distracter, appeared at extrafoveal locations on the screen, and they were rewarded for detecting a subtle color change in the target. Before the presentation of the gratings, distracter images appeared briefly at locations between the fixation stimulus and the target and distracter gratings.

First we examined if there were any differences in performance between trials with images and trials without distractor images (Fig. 2A). Sensitivity $d'$ was lower in trials with distractor images than in trials with no images [$t(5) = -4.73, P < 0.01$]. RT in trials with images was faster than RT in trials with no images [Kruskal–Wallis $H(1) = 8.68, P < 0.005$]. To control for the possibility of a speed-accuracy tradeoff, we applied the “EZ-diffusion model” (31) (Materials and Methods), which uses accuracy and RT data and expresses performance in terms of underlying variables neutral to the speed-accuracy tradeoff: drift rate, boundary separation, and nondecision time. Drift rate is closest to a combination of reaction time and accuracy. In our data, the presence of images reduced drift rate, confirming the reduction in $d'$ [$t(5) = -5.52, P < 0.005$].

There was a significant effect of stimulus onset asynchrony (SOA) on $d'$ [ANOVA $F(2,17) = 15.4, P < 0.001$], as shown in Fig. 2B. Post hoc comparisons revealed that $d'$ was lower in the longest SOA (500 ms) than in shorter SOAs and lower than in trials with no distractor image (all $P < 0.05$). Thus, distractors that appeared the longest time before the target change attracted attention most effectively. There was also an effect of SOA on RT [Kruskal–Wallis $H(2) = 583.1, P < 0.001$] with RT decreasing as a function of SOA. RT was faster when there was a distractor image than when there was no image, even at SOA 50 ms. We suspect that image onset functioned as a cue to “get ready” for the target change, thus causing RT to decrease. The drift rate confirmed that SOA negatively affected the ability to detect the change [ANOVA $F(2,17) = 5.92, P < 0.05$] and therefore unlikely the result of mere speed-accuracy tradeoff. Further evidence that speed-accuracy tradeoff was not a factor is seen in Fig. 2B (Left and Center), showing that sensitivity was reduced only at the longest SOA (500 ms), whereas RT decreased systematically from short to long SOAs.

The variations in SOA were confounded with image duration (distractor images remained on the screen until the monkeys responded). Therefore, in a control experiment, we tested whether it was onset time or image duration that was the relevant factor. In control sessions in two monkeys (L and P), images were presented for a constant duration of 50 ms, but varied onset times. If onset time were the relevant factor, $d'$ should decrease as SOA gets longer, just like in the main task. The result (Fig. 2C) shows that this was not the case. Thus, duration of the distractor images, not their timing, was the relevant factor.

In this control experiment, $d'$ in general, including in the no-image condition, was lower than in the main experiment. One possible explanation is that the temporal structure of the task may set the monkeys’ expectations about when the target is most likely to change and that a change in temporal structure reduces general performance, affecting all conditions to about the same extent. Thus, the comparison between conditions seems valid.

In the main experiment, we found an interaction between SOA and trial length, which is the time period from the onset of the gratings until the color change of the target grating (Fig. 2D) [ANOVA interaction SOA × length $F(4,45) = 18.9, P < 0.0005$]. The long SOA resulted in the lowest $d'$ regardless of trial length, but the difference between short and long SOAs was larger in long trials than in short trials, indicating that the monkeys became more distracted as the trial got longer.

We separated trials based on whether the face image was on the same side as the target location (congruent) or not (incongruent). For $d'$, there was a significant interaction between congruence and emotional valence in distracter faces [ANOVA interaction $F(1,20) = 5.65, P < 0.03$]. Among trials with emotional faces, $d'$ in congruent trials was lower than in incongruent trials [$t(5) = -3.93, P < 0.02$], whereas with neutral faces, the difference
was not significant \( t(5) = 1.97, P = 0.11 \), suggesting that spatial location partly determines the saliency of emotional distracters. Conversely, emotional distracters become easier to ignore when they are located in the opposite hemifield from the target.

The animals made a saccade toward the target while the images were still on the screen. Thus, in the congruent trials, the saccade crossed a face image to reach the target, whereas in incongruent trials, the saccade crossed a scrambled image (as illustrated in Fig. 3A). We examined whether facial expression affected saccadic eye movements. Saccade end points and saccade trajectories did not vary with distracter face expressions (Fig. S1). The peak saccade velocity varied with expression and congruence in the 500-ms SOA trials (Fig. 3B). There was a significant interaction between congruence and emotional expression in trials with 500-ms SOA [rank aligned transform/ANOVA interaction \( F(2,359) = 3.42, P < 0.05 \)]. For threat faces, peak saccade velocity was significantly higher when congruent than when incongruent with the target. The same comparisons in the fear and neutral categories were not significant. Thus, OT inhalation appeared to reduce the distraction caused by threat faces.

The foregoing results indicate that faces with affective content are potent distracters in our task. To test whether the interference can be reversibly affected by manipulating brain circuitry, the monkeys were treated with the hormone OT, given previous reports that it has a specific effect on the amygdala and forebrain circuits and influences social behavior (32, 33).

Fig. 4A shows the results of OT inhalation on d' and RT when pooled across SOA and facial expression. In trials with no image, d' increased significantly on OT compared with baseline \( t(5) = -2.68, P < 0.05 \). Although d' among trials with images did not change in general, there was a dependency on SOA and emotional expression (Fig. 4B). In the longest SOA (500 ms), especially in trials with threat faces, d' increased on OT [ANOVA interaction SOA × treatment × expression \( F(2,87) = 3.85, P = 0.025 \)]. Between baseline and OT, trials with threat faces at SOA 500 ms were significantly different \( t(10) = -2.41, P < 0.05 \), whereas the same comparisons for categories fear and neutral were not significant [neutral: \( t(10) = -0.32, P = 0.75 \); fear: \( t(10) = -0.46, P = 0.65 \)]. Thus, OT inhalation appeared to reduce the distraction caused by threat faces.

Treatment with OT also significantly increased RT in trials with face distracters regardless of expression (Wilcoxon rank sum test \( z = -5.09, P < 0.000001 \)). Importantly, the change in RT in trials with no images was not significant when Bonferroni-corrected for multiple comparisons (34) (Wilcoxon rank sum test \( z = 2.01, P = 0.04 \); Fig. 4A), indicating specificity of the OT effect with regard to face distracters. The interaction between treatment and presence of a distractor image was significant [aligned rank transform/ANOVA \( F(1,4189) = 10.63; P < 0.005 \)]. The RT increase was not dependent on SOA or expression (Fig. 4C). Furthermore, the effect of target/face congruence on d' in the baseline condition appeared to be reduced on OT (Fig. 4D), although this effect was not significant.

The effect of congruence on saccade velocity seen in the baseline was no longer present after OT treatment. This effect is also indicated by a significant interaction between treatment and congruence [ANOVA interaction treatment × congruence \( F(1,787) = 4.06, P < 0.05 \). Fig. 5A shows the combined the
results of the congruence effect in baseline and OT conditions. Mean saccade velocity relative to saccade landing time is in Fig. S2.

There was no effect of expression on pupil response in the OT condition [ANOVA F(2,141) = 0.35, P = 0.69] in contrast to the baseline condition, where the pupil response to threatening faces was smaller than to neutral and fearful faces (Fig. 5B).

Finally, we estimated the latency of the OT effect by calculating a moving average of d‘ over time (Materials and Methods). The progression is shown in Fig. 5C. The effect latencies were 72 and 66 min for d‘ and RT, respectively.

Discussion
We examined the effect of irrelevant emotional faces on monkeys’ performance of an attention task, which required detecting a subtle color change in one of two gratings. The distracter faces were interposed between the test stimuli and the fixation spot, and the monkeys never directly looked at the face images. Faces interfered with sensitivity to detect the color change, especially when the faces had an emotional expression. The effect of expression was only observed at the longest presentation durations, suggesting that the faces did not trigger an immediate shift of attention. Faces with a threat expression affected saccade velocity when the monkeys made the eye movement in response to the color change in the target. Threat faces were also associated with smaller pupil constriction compared with faces with fear and neutral expressions. Application of intranasal OT mitigated the effects of emotional faces on sensitivity for detecting the color change, saccade velocity, and pupil response while increasing reaction times overall.

Our results confirm that rhesus monkeys have an attentional bias toward faces, emotional faces in particular, similar to humans (2, 30, 35, 36). As a result, the presence of faces can weaken the processing of concurrent stimuli (25, 37). The stronger effect for threat faces supports Ohman’s threat advantage hypothesis (38), and corroborates extensive work in humans showing that affectively significant stimuli can influence behavior even when irrelevant to the task at hand.

One common hypothesis is that emotional stimuli are processed through a rapid preattentive subcortical pathway (39). Through this pathway, a strong enough signal might dominate the saliency map and thus capture attention (10, 40). However, rapid attentional capture by briefly presented emotional stimuli, as found in many human studies (11–13, 37), was not observed in our experiment. Because attention is considered to be resource limited, one explanation may be that the load imposed in the attention task was high. The likelihood of attention capture has been shown to be greater under low load conditions, when residual capacity was available, than under high load conditions (14, 41, 42). We titrated the difficulty of the target color change so the monkeys’ performance was between 70% and 90% correct, creating a high attentional load. Thus, our results are consistent with the idea that the processing of affective stimuli is gated by attention (16). This conjecture needs further evaluation by systematically varying load in future experiments.

Saccade velocity was influenced by facial expressions. Although saccade velocity is not under voluntary control (43), it is known to increase with arousal (44). Viewing a threatening face may increase arousal and therefore increase saccade velocity. We have examined saccade trajectories but found no effects. Threat faces also produced an autonomic response as evidenced by a reduced pupillary light reflex in response to image onset. Pupil diameter may increase in proportion to mental effort (45), arousal (46), and attention (47). In humans, threat of shock reduced the pupillary light reflex (48), whereas diazepam antagonized the effect (49). Thus, the effect we observed may be due to a task-related expectancy of seeing threatening faces. Stimulation of the amygdala results in pupil dilation (49), possibly through interconnections with the locus coeruleus; therefore, this effect could be the result of activity in the subcortical pathway.

Administration of OT reduced the effects of emotional faces on the primary task. The latency of our behavioral effect of OT (~70 min) is in line with previously reported behavioral effects (peaking at 110 min) using the same nebulization method (50). CSF measurements (50) and microdialysis in the amygdala and hippocampus (51) indicate increased levels of OT 30–60 min after delivery. The amygdala is further implicated by rodent studies showing that axonal release from oxytocin-positive neurons from hypothalamic nuclei in the central amygdala reduces fear responses (52). Like diazepam, OT acts on components of the GABAergic circuit in the central amygdalar complex (53). However, other brain structures may be involved as well. Rapid detection of emotional stimuli takes place even when the amygdala is lesioned (54–56). In humans (57) and in monkeys (58), OT has been shown to reduce responses to negative facial expressions not only in the amygdala, but also in inferior temporal cortex and prefrontal cortex (58). Therefore, the perceptual advantage of emotional stimuli and the dampening of the effect by OT likely involve a brain network including those areas.

Materials and Methods
Animals. All procedures were in accordance with the National Institutes of Health and US Department of Agriculture guidelines and approved by the Massachusetts Institute of Technology Committee on Animal Care. Three macaques (Macaca mulatta) named L, P, and H were used. Each animal was surgically implanted with a head post before training. Surgery was conducted under aseptic conditions with isoflurane anesthesia, and antibiotics and analgesics were administered postoperatively.

Tasks. Stimuli were presented on an LCD monitor (resolution 1,280 × 768 pixels, gamma corrected) at a distance of 57 cm. Presentation of stimuli and behavioral parameters were controlled using PsychToolBox and Eyelink Toolbox (59, 60) on Matlab software. Eye position was detected by an infrared based eye-tracking system (1,000-Hz Eyelink; www.s-research.com) and recorded using a Plexon MAP data acquisition system. The animals were rewarded with fruit juice or water.

The monkeys performed a covert attention task. Each trial started with a white fixation spot of 0.4 × 0.4° in the center of the screen for 500 ms, and monkeys were required to acquire fixation within this period. The monkeys had to hold their gaze within a 1–1.5° square window centered on the
fixation spot throughout the trial or the trial was aborted. After 500 ms, two colored, vertical, square wave gratings of 1.5° diameter appeared at 10° to the left and right of the fixation spot. One grating was red and the other blue, randomly assigned each trial. The moment the gratings appeared, the color of the fixation spot changed to a color that matched one of the gratings. The change in fixation spot color was the cue for the animal to monitor the matching grating (the target) for a subtle color change. The monkeys were rewarded for making a saccade to the target grating, which counted as a correct response. The other grating (the distractor) could also be correct but was not rewarded, and the trial was ended and counted as an error. If the monkeys did not make a saccade to either grating within 800 ms after the target change, the trial was aborted without reward and counted as an error.

The timing of target and distractor changes was based on independent random picks between three possible change times for each grating: 600, 900, and 1500 ms. In 35% of the trials, only the target changed. Monkeys L and P were highly trained on this task, whereas monkey H had no previous exposure to similar tasks and reached required proficiency after several weeks of training. The difficulty of the color change was titrated such that the monkeys scored 70–90% of the trials correct. The CIELab color space coordinates of the gratings were as follows: red prechange, 45.4, 78.8, 80.0; postchange (min), 47.5, 64.9, 50.6; postchange (max), 49.7, 70.5, 61.7; blue prechange, 27.3, 51.3, 102.7; postchange (min), 21.3, 50.1, 90.7; postchange (max), 24.1, 53.7, 99.4.

In 90% of the trials, unscrambled or scrambled photographs of monkey faces appeared on the screen at 50, 200, or 500 ms before the target change, as illustrated in Fig. 1 A and B. The image remained on the screen until the monkey responded to the target change or broke fixation. The size of the images was 8° × 8°, and they were centered at 5° eccentricity, between the fixation spot and the gratings. Image size and location was varied in a separate set of sessions (size varied between 5° and 8° width and location was varied in the horizontal dimension ±4° from midline), with no obvious effect on the pattern of results. The images were irrelevant to solving the task. Randomization of the identity and location of the images ensured that they were not predictive of the target change. Because trials were immediately aborted if the monkeys broke fixation, they did not get the chance to directly look at the images. Initially the monkeys were not capable of doing the task with the images at full brightness, presumably because they were too distracting or because the luminance contrast of the images reduced perceptual saliency of the target gratings. Both monkeys had seven sessions during which the brightness of the images was gradually increased until they were at full brightness. Here we only include data from sessions after this initial training. Using the luminance profile of the red, green, and blue guns of the monitor, measured using a Colorvision Spyder 2 Pro photometer, the mean luminance of each image was calculated, and the brightness of each image was adjusted to make mean luminance of all images 21 Cd/m². Background luminance was 4.18 Cd/m².

Two images appeared simultaneously, one on each side of the fixation spot. One was a color headshot of monkey, and the other was a scrambled version of the image. Monkey faces were from a database created in the laboratory of David Amaral (University of California, Davis, Sacramento, CA) and used with permission. Each headshot had one of three facial expression categories. Neutral expressions were labeled neutral, open mouth threat expressions were labeled threat, and bare teeth grin and lip smacking expressions were combined in the category fear. Because of this combination, the label fear does not always describe the emotion being expressed, because lip smacking signals intent to engage in affiliation rather than fear (61). However, when analyzed separately, the two expressions yielded qualitatively similar results: 28% of images in the category fear were bare teeth grin. We only used individuals for which we had at least one of image of each category. The stimulus set was split into subsets where each subset contained four to six individuals. The subset used was based on a session by session basis. Two sessions per day were done (one baseline session and one treatment session).

Hormone/Saline Administration. To examine the effect of the hormone OT on performance of the task, it was administered intranasally using a protocol described in Chang et al. (50). After fixing the head using the head post, the experimenter applied a silicone mask connected to a Pari (www.pari.com/products/nebulizers.html) baby nebulizer fully covering the mouth and nose. A foam lining was applied to the mask to minimize leakage. Either saline or OT (25 IU/mL; Agrilabs) was delivered via nebulization continuously for 5–15 min. Before experimental sessions, monkeys were first be habituated to the nebulizer procedure involving placement of a mask and saline delivery using the nebulizer in an incremental fashion until they seemed comfortable during the procedure. Habituation took about 1 wk. Once monkeys were habituated, the nebulizer procedure, testing began. On each day of testing, monkeys were given OT 60 min before doing the task. Each monkey performed two sessions. Monkey P had 720 baseline trials and 729 on OT, monkey L had 786 baseline trials and 744 on OT, and monkey H had 1,054 baseline sessions and 766 on OT. We examined the effect of saline inhalation in one animal and compared it with sessions without the inhalation procedure (baseline). There was no significant difference between baseline and saline conditions.

Analysis. The performance measures included RT, sensitivity (d'), peak saccade velocity, and pupil diameter. Because the monkeys responded by making an eye movement to one of the gratings, RT was defined as the time between the target change and the gaze entering a radius of 5° around target or distractor grating. Outliers in RT were detected and removed using iterative implementation of the Grubbs test (62). As RT data were not normally distributed, hypothesis testing was done using the Kruskal-Wallis test. To test for interactions, the data were aligned rank transformed, followed by ANOVA (63).

The sensitivity measure d' was calculated by subtraction of the Z-score for the false alarm (FA) rate from the Z-score of the hit (H) rate. FAs were considered to be trials in which the monkeys made a saccade to the target grating before the target had changed. Responses to distractors were very uncommon (typically <1%). Hypothesis testing on d' was performed using ANOVA for repeated measures.

Per-millisecond saccade velocity was calculated as the derivative of the vector magnitude created by the x and y eye movement channels. The peak velocity was the maximum velocity between the target change and the eye entering the specified radius following the saccade response to the grating.

Speed-accuracy tradeoff can make it difficult to draw conclusions about whether the subject’s ability is affected by the experimental manipulation (64). One way to deal with this is to use a model to estimate a set of variables that underlie performance and not subject to speed-accuracy tradeoffs. Ratcliff’s diffusion model (65) is one such model. Here we apply a simplified version, the EZ diffusion model (31). The simplified model takes as input for each condition: mean RT, variance of RT, and percent correct. Output are three underlying variables: drift rate, boundary separation, and nondecision time. Drift rate, which we report, is closest to a measure combining speed and accuracy.

The latency of the OT effect was estimated by calculating a moving average in a sliding window stepping through the session in one-trial steps. The sliding window was 200 trials wide, and the first trial in the window marked time; t tests were done at each step. When three consecutive steps were significant, the first was taken as the start of the effect, in number of trials. To get the latency in seconds, this was multiplied by mean trial duration across sessions including intertrial intervals and idle periods, which amounted to 8 s. The 60-min wait time between OT administrations and testing was added to yield the latency.

ACKNOWLEDGMENTS. We thank Dr. David Amaral for kindly providing us with the stimuli used, Dr. David Osher for helpful discussions, and Emma Myers for training the animals and setting up the training rig. This research was supported in part by grants from The Simons Foundation (to M.S.) and National Institutes of Health Grant EY017292 (to R.D.).