

# Neuronal population correlates of target selection and distractor filtering

Elaine Astrand, Claire Wardak, S. Ben Hamed

▶ **To cite this version:**

Elaine Astrand, Claire Wardak, S. Ben Hamed. Neuronal population correlates of target selection and distractor filtering. NeuroImage, 2020, 209, 10.1016/j.neuroimage.2020.116517 . hal-03045183

**HAL Id: hal-03045183**
**https://hal.science/hal-03045183**

Submitted on 7 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NeuroImage 209 (2020) 116517

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

# Neuronal population correlates of target selection and distractor filtering

Elaine Astrand [a,b,**], Claire Wardak [a,c], Suliann Ben Hamed [a,*]

[a] Institut des Sciences Cognitives Marc Jeannerod, Département de Neuroscience Cognitive, CNRS UMR 5229, Université Claude Bernard Lyon I, 67 Boulevard Pinel, 69675, Bron Cedex, France
[b] School of Innovation, Design, and Engineering, Mälardalen University, Högskoleplan 1, 721 23, Västerås, Sweden
[c] Imagerie et Cerveau, INSERM U1253, Université de Tours, Faculté de Médecine, 10 Boulevard Tonnellé, 37032, Tours Cedex 1, France

ARTICLE INFO

*Keywords:*
Frontal eye fields
Neuronal population
Perception
Decoding
Target selection
Distractor filtering
Spatial processing

ABSTRACT

Frontal Eye Field (FEF) neurons discriminate between relevant and irrelevant visual stimuli and their response magnitude predicts conscious perception. How this is reflected in the spatial representation of a visual stimulus at the neuronal population level is unknown. We recorded neuronal population activity in the FEF while monkeys were performing a forced choice cued detection task with identical target and distractor stimuli. We quantified, using machine learning techniques, estimates of target and distractor location from FEF population multiunit activities. We found that the FEF population activity provides a precise single trial estimate of reported stimuli locations. Importantly, the closer this prefrontal population single trial estimate is to the veridical stimulus location, the higher the probability that the target or the distractor is reported as perceived. We show that stimulus perception is rescued by the estimate of attention allocation specifically when the latter is close enough to the actual stimulus location, thus indicating a partial independence between attention and perception. Overall, we thus show that how and what we perceive of our environment depends on the spatial precision with which this environment is coded by prefrontal neuronal populations.

## 1. Introduction

Perception is often defined as the ability to become aware of one's environment through the senses. How we perceive our surroundings is influenced both by internal voluntary top-down processes, whereby higher priority is given to relevant aspects of the environment relative to irrelevant aspects, and by external involuntary bottom-up processes, whereby intrinsically salient items impose themselves onto our perception. The outcome of perception may be veridical, whereby reality is perceived consciously, or it may be erroneous, misrepresenting parts of reality or not reaching consciousness.

The dorsolateral prefrontal cortex, including the frontal eye fields (FEF), is proposed to play a key role in conscious perception (Thompson and Schall, 1999, 2000; Libedinsky and Livingstone, 2011; Panagiotaropoulos et al., 2012; Vugt et al., 2018). Specifically, the magnitude of FEF neuronal responses to a target stimulus has been shown to correlate with the behavioral outcome, with higher magnitude during hit and false alarm trials as compared to miss and correct rejection trials, respectively (Thompson and Schall, 1999, 2000). These results point towards the involvement of FEF in generating the contents of visual perception that subsequently lead to the production of overt behavioral report. This is further supported by evidence from electrical micro-stimulation of the macaque FEF leading to enhanced perception (Moore and Fallah, 2004) and inactivation of the FEF that induces deficits in visual search target detection independent of search difficulty (Wardak et al., 2006).

However, perception does not only involve selecting a target, but also filtering out competing distractors. It has long been established that FEF activity demonstrates this ability. Specifically, in pop-out visual search tasks, i.e. in tasks in which the target can easily be distinguished from the distractors, neuronal responses to a distractor presented in the neuron's receptive field are suppressed (Schall and Hanes, 1993; Thompson et al., 1996). As the distractor becomes more similar to the target (in more difficult visual search tasks), the neuronal activity in the FEF becomes less reliable at discriminating the target from the distractor (Sato et al., 2003). Furthermore, the degree of distractor suppression as assessed from overt behavior has been shown to correlate with the degree of neuronal suppression in the FEF and inactivation of the prefrontal cortex induces a significant increase in distractibility, i.e. in the production of undesired responses to intervening distractors (Suzuki and Gottlieb, 2013). The erroneous selection of a distracter in the FEF has been shown

* Corresponding author.
** Corresponding author. School of Innovation, Design, and Engineering, Mälardalen University, Högskoleplan 1, 721 23, Västerås, Sweden.
*E-mail addresses:* elaine.astrand@mdh.se (E. Astrand), benhamed@isc.cnrs.fr (S. Ben Hamed).

https://doi.org/10.1016/j.neuroimage.2020.116517
Received 16 September 2019; Received in revised form 5 December 2019; Accepted 1 January 2020
Available online 8 January 2020
1053-8119/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

to correlate to errant allocation of covert attention prior to the saccade (Heitz et al., 2010), suggesting that attentional orienting guides stimulus selection.

In addition to FEF's involvement in the selection of relevant stimuli and the construction of perception, neurons in the FEF are also shown to be selective to the spatial location of a target stimulus (Thompson et al., 2005). By comparing the latency of FEF spatially selective responses in Local Field Potentials (LFPs) and single-unit spiking activity, results from Monosov et al. (2008) indicate that the spatial selection of a visual stimulus is locally computed in the FEF most probably by using fast non-spatially selective activity relayed from earlier visual areas. In a cued visual search task, Monosov and Thompson (2009) link spatial selectivity in the FEF to visual perception by showing that the magnitude of the spatially selective neuronal response is significantly correlated with the accuracy and speed of target identification (Monosov and Thompson, 2009). In this study, monkeys had to use spatial information relayed by the cue and the target but were rewarded for reporting target identity. The authors argue that the spatially selective activity in the FEF plays a causal role in the enhancement of visual object-related perceptual processing and is associated with spatial attention. This study, using single-unit neuronal activity, is further corroborated by showing that simultaneously recorded multi-unity activities from the FEF (and lateral prefrontal cortex) encode covert spatial attention and is predictive of behavior (Tremblay et al., 2015; Astrand et al., 2016). However, it remains unknown whether the spatially selective neuronal activity in the FEF after target presentation, is associated with perception (i.e. perceived stimulus location) or merely reflects reorienting or reinforcement of spatial attention.

The prefrontal neural correlates of perception and distractor filtering have most often been studied from the perspective of the response of single neurons. Cohen et al. (2010) investigated neuronal functional interactions during target perception. They show specific functional cooperative interactions between neurons with overlapping receptive fields and competing interactions between neurons with non-overlapping receptive fields. This suggests complex integrative mechanisms at the level of the population but doesn't provide a quantification of population information. Some studies have investigated the impact that a neural population has on perception by artificially reconstructing a population from single neurons recorded independently (Bichot et al., 2001; Astrand et al., 2015). While this approach does allow to capture stable information across trial population coding schemas, single trial population information is lost.

To address whether perception is spatially encoded in the FEF neuronal population, we use simultaneously recorded Multi-Unit Activity (MUA) and apply machine learning techniques to identify how targets and distractors are represented by the FEF population on a given trial and how this representation is predictive of overt behavior. We use a forced choice cued target detection task, in which trained monkeys are required to respond as fast as possible to a low saliency target presented at a cued location while at the same time ignoring distractors identical in all respects to the target and presented at other uncued locations. We show that when a stimulus is reported, whether this stimulus is the target of behavior or a distractor, the neuronal FEF population precisely encodes its location. In other words, the population constructs an accurate representation of the stimulus in space. In contrast, when a stimulus is not reported, whether this stimulus is the target of behavior or a distractor, its spatial representation, as encoded by the FEF, does not match its real location. We describe a strong correlation between the error in the estimation of the position of the visual stimulus in space as coded by the neuronal population with respect to its actual physical location, and overt behavior. Finally, we further show that the spatial estimate of the visual stimulus is partly independent from the spatial estimate of the cued location prior to the stimulus onset, and that each monkey relies differently upon the spatial encoding of the cue and target in their behavioral optimization. Overall, we propose that the visual percept of a reported stimulus is not only accounted for by the strength of the visual representation but also by how accurately its spatial location is encoded in the neuronal population.

## 2. Methods

### 2.1. Surgical procedure and FEF mapping

All experimental procedures were identical to those used in Astrand et al. (2016). One head fixation post and two MRI compatible PEEK recording chambers were placed over the FEF, one in the left and one in the right hemispheres of two male rhesus monkeys (Macaca mulatta) weighing between 6 and 8 kg. During the surgery, gas anesthesia was provided to monkeys using Vet-Flurane, 0.5–2% (Isofluranum 100% at 1000 mg/g) followed by an induction with Zolétil 100 (Tiletamine at 50 mg/ml, 15 mg/kg and Zolazepam, at 50 mg/ml, 15 mg/kg). Post-surgery pain was controlled with a Morphine pain-killer (Buprecare, buprenorphine at 0.3 mg/ml, 0.01 mg/kg), 3 injections at 6 h interval (first injection at the beginning of the surgery) was administered post-surgey and a full antibiotic coverage was provided with Baytril 5% (a long action large spectrum antibiotic, Enrofloxacin 0.5 mg/ml) at 2.5 mg/kg, one injection during the surgery and thereafter one each day during 10 days. In order to have a precise localization of the arcuate sulcus and surrounding gray matter underneath each of the recording chambers, a 0.6 mm isomorphic anatomical MRI scan was acquired post surgically on a 1.5T Siemens Sonata MRI scanner, while a high-contrast oil-filled 1mmx1mm grid was placed in each recording chamber, in the same orientation as the final recording grid. The FEF was defined as the anterior bank of the arcuate sulcus and sites were specifically targeted in which a significant visual and/or oculomotor activity was observed at 10° to 15° of eccentricity from the fixation point during a memory guided saccade task. In order to maximize task-related neuronal information at each of the 24-contacts of the recording probes, we only recorded from sites with task-related activity observed continuously over at least 3 mm of depth. All procedures were approved by the local animal care committee (C2EA42-13-02-0401-01) in compliance with the European Community Council, Directive 2010/63/UE on Animal Care.

### 2.2. Behavioral task

A 100% validity cued luminance change detection task with temporal distractors (Fig. 1A) was used. With their head fixed, the monkeys were placed in front of a computer screen (1920x1200 pixels and a refresh rate of 60 Hz). To initiate a trial, they had to hold a bar in front of the animal chair, thus interrupting an infrared beam. On trial initiation, a blue fixation cross (0.7 × 0.7°) appeared in the center of the screen and the monkeys were required to hold fixation throughout the entire trial until the manual response, within a fixation window of 2° × 2°. Break of fixation aborted the trial. After correct manual responses, both monkeys tended to produce a saccade towards the reward dispenser.

Four gray square landmarks (0.5° × 0.5° for monkey M1, 0.68° × 0.68° for monkey M2) were presented simultaneously with the fixation cross and were placed at an equal distance from the fixation point, in the upper right, upper left, lower left and lower right quadrants of the screen, thus defining the corners of an imaginary square. To ensure that the recorded neurons represented the cued spatial location, we adjusted the eccentricity of the landmarks from day to day between 10° and 15°, as inferred from the neurons' response to a memory-guided saccade task with saccadic targets placed at variable locations in this range. After a variable delay from fixation onset, ranging between 700 and 1900 ms, a green squared cue was presented for 350 ms, indicating to the monkey in which of the four landmarks the rewarding target change in luminosity would take place. The cue was small (0.2° × 0.2° for monkey M1 and 0.3° × 0.3° for monkey M2) and it was presented close to the fixation cross in the same direction as the landmark to be attended (at 0.3° for monkey M1 and at 1.1° for monkey M2, from the fixation point). As a result, the cued target detection task was more difficult than if the cue

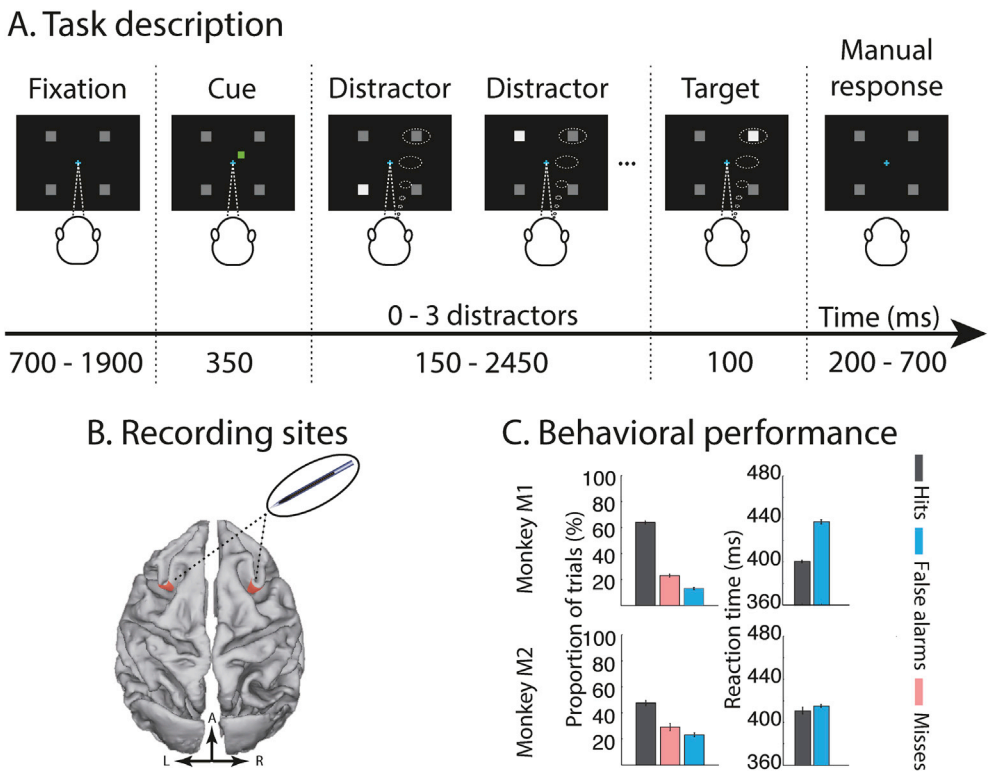



**Fig. 1.** (A) **Task description.** A trial was initiated by the simultaneous onset of a fixation point and 4 Gy landmarks. Monkeys were required to hold a bar and fixate the fixation point throughout the trial. After a variable delay ranging from 700 to 1900 ms, a small green square was presented near the fixation point, indicating the location in which the target will be presented. During a variable delay, ranging from 500 to 2800 ms, monkeys were required to orient their attention to the cued landmark to detect a small change in luminosity. During the delay, a change in luminosity could occur on any of the other three landmarks and monkeys were required to ignore them. A liquid reward was distributed to the monkeys for releasing the bar 200–700 ms after target luminosity change. (B) **Recording sites.** On each session, two 24-contact recording probes were placed, one in each FEF. (C) **Behavioral performance.** Median across sessions (n = 15) of the proportion of hit trials (dark gray bars), misses (light gray bars), and false alarms (intermediate gray bars) are depicted for each monkey separately. Error bars correspond to median absolute deviation across sessions.

had been presented at expected target location, thus resulting in more error trials. After cue presentation, the monkeys needed to orient their attention to the target landmark in order to monitor it for a brief change in luminosity (100 ms) while maintaining eye fixation on the central cross. The change of luminosity was set independently for each position to generate a near-threshold perceptual change for each target location, as estimated from the behavioral performance. The change in target luminosity could occur anywhere between 500 and 2800 ms from cue onset according to a uniform probability distribution. In order to receive a water or juice reward, the monkeys were required to release the bar (thus restoring the infrared beam) in a time window of 200–700 ms following the change in target luminosity (hit trial). In order to make sure that the monkeys were correctly orienting their attention towards the cued landmark, unpredictable changes in the luminosity, identical to the awaited target luminosity change, could take place at the non-cued landmarks (distractors). On each trial, from none to three such unpredictable distractor luminosity changes could take place, no more than one per non-cued landmark position. The monkeys had to ignore these distractors. Responding to such a distractor interrupted the trial and was counted as a false alarm trial if the response fell within 200–700 ms following the distractor. Failing to respond to the target (miss trial) similarly aborted the ongoing trial.

### 2.3. Neural recordings

Bilateral simultaneous recordings in the two FEF hemispheres were carried out using two 24-contact Plexon U-probes. The contacts had an interspacing distance of 250 μm. Neural data was acquired with the Plexon Omniplex® neuronal data acquisition system. The data was amplified 100 times and digitized at 40,000 Hz. The neuronal data was high-pass filtered at 300 Hz. In the present paper, all analyses are performed on the multi-unit activity recorded on each of the 48 recording contacts. A threshold defining the multi-unit activity was applied independently for each recording contact and before the actual task-related recordings started. All further analyses of the data were performed in Matlab.

### 2.4. Discrete classification procedure

Decoding analyses were performed in Matlab. A regularized linear regression was used to investigate whether the neural population contained information about target and distracter location (four possible locations). A linear regression defines the weight matrix W that minimizes the mean square error of C=W*(R + b), where C is the class (here, the spatial position, amongst four possible locations), b is the bias and R is the neural response (here, a 48 element vector representing the neuronal multi-unit activity at each of the 48 recording contacts, at the time of interest; for each recording channel and each trial). The multi-unit activity was smoothed by averaging the spiking activity over 150 ms sliding windows (resolution of 1 ms); this window width corresponds to a trade-off between decoding performance and decoding speed, as narrower filtering windows result in a lower performance while wider filtering windows decrease temporal resolution (Farbod Kia et al., 2011). To avoid over-fitting we used a Tikhonov regularization which gives us the following minimization equation: norm(**W***(**R** + **b**) – **C**)+ λ*norm(**W**). The scaling factor λ was chosen to allow for a good compromise between learning and generalization (Astrand et al., 2014). Specifically, the decoder was constructed using two independent regularized linear regressions, one classifying the x-axis (two possible classes: −1 or 1) and one classifying the y-axis (two possible classes: −1 or 1). Each regression used the following procedure. The neural data was divided into training (70% of data, M1: average of 161 trials per session and M2: average of 103 trials per session) and testing set (30% of data). The training set was used to define the weight matrix, W and the testing set was then multiplied by W, and the sign was evaluated to yield the predicted class C, for these novel trials (−1 or 1). Final classification performance was calculated by dividing the number of correct predictions of the classifier on test trials by the overall testing sample size. As a result, the classification performance is a measure, ranging between 0 and 100%, of how well information in the neural population allows to predict the quadrant in which the stimulus of interest is being presented (i.e. target or distracter), at a specific time in the trial. When the train- and test-sets of neuronal activities correspond to the same timing relative

to the key events of the task, the classification performance is a measure of the instantaneous information content of the variable of interest at exactly that time.

### 2.5. Two-tailed non-parametric random permutation

Due to task configuration, absolute chance level is at 25%. However, in order to define the statistical significance of the reported classification performance, we defined for each classifier, the 95% confidence interval limit as follows. For each recording channel, we reassigned random labels (e.g. relative to the cue) to each trial and performed the same classification analysis as described above. This procedure was repeated 1000 times and yielded a 1000 data point distribution of chance classification performance. Classification performance for real non-permuted data was considered significantly above or below chance if it fell within the 2.5% upper or lower tail of this random permutation distribution (2-tailed non-parametric random permutation test, 0.05 alpha level).

### 2.6. Continuous classification procedure

As a complement to the above discrete classification procedure, we also investigated the continuous two-dimensional (x,y) output of the classifier that provides a more precise localization of the spatial variable of interest, e.g. cued or target location. The training procedure is identical to that described above in that it trains the classifier to associate the data to the cued (x,y) position, where x and y only takes −1 and 1 values. The testing procedure is slightly different in that it tests novel data without assigning the output of the classifier to a class (−1 or 1). Instead the continuous output from the decoder is interpreted as coordinates in the x/y-plane, i.e. a (x,y) position, that reflect the error to the expected stimulus location. The result of this analysis can thus be read as the spatial locus of the variable of interest at any point in time. The distance between this decoded spatial position (of for example perceived target location) and the actual physical location (of the target) is calculated and hereafter termed “target to decoded target distance”. The decoded spatial locus of the perceived target can also be represented as a spatial probability map, constructed by calculating the statistical significance of a given (x,y) classifier output at each spatial location using non-parametric random permutation test (see previous section) and visualizing these as a color coded z-scores. Specifically, a map of p-values was constructed by comparing the actual data with randomly permuted data. Z-scores were then calculated from these p-values yielding the signed number of standard deviations from the normal mean probability, at each spatial location, that is significantly over- or under-represented ($p < 0.01$).

### 2.7. Euclidian multidimensional distance

The multidimensional Euclidian distance (hereon named MDD) between multiunit activity for all stimulus position-pairs in one session can be considered to reflect the spatial selectivity of the recorded neuronal population to the four target/distractor positions. This measure represents how distant the population neuronal responses to each target sit in the higher population multidimensional space. The furthest away, the larger the MDD. It is to be noted that MDD is calculated using the averaged magnitude of multiunit activity across all trials for each target/distractor position (for each channel), meaning that reliability in the neuronal response is neglected in favor of the magnitude, contrasting with the decoding approach that depends on both measures. The MDD was calculated per trial as follows:

$$\mathrm{MDD}(\mathrm{pos}_1, \mathrm{pos}_2) = \sqrt{\sum_{i=1}^{48} \left(r_i^{pos1} - r_i^{pos2}\right)^2}$$

where r corresponds to the averaged neuronal response across trials (of a specified target position) for each channel i. The MDD for all position-pairs (six pairs) were averaged to yield one MDD-value per session which can be considered as a measure of the four-dimensional spatial selectivity (four positions). The MDD per session is therefore in units of MUA/s and represents the multi-dimensional MUA difference between the four different target/distracter locations.

### 2.8. Relationship between target position population decoding accuracy and single channel response statistics

All through this work, we will be quantifying the available information about the cued and target/distractor location in the prefrontal FEF neuronal ensembles being recorded from, using a regularized linear regression procedure that associates bilateral neural FEF response patterns at a given time from target presentation with the location of the expected target. An important question is how this decoding accuracy relates to the single channel response statistics. For each session, we computed a measure of spatial selectivity that recapitulates single channel response statistics (reflecting mainly the magnitude of MUA) in the high-dimensional recording space and investigated its relation to the decoding accuracy (Fig. 2C). We calculated the Euclidian multi-dimensional distance (MDD) between average MUA responses 100–200 ms after target presentation, for each of the six possible pairs of target locations. The average of these 6 MDDs was computed for each session. This measure reflects the average separation in neuronal responses to distinct target locations in the high-dimensional recording space (i.e. spatial selectivity). This measure captures the discrimination capabilities of the neuronal population based on mean individual channel MUA responses. The MDD is highly correlated with decoding accuracy across sessions (Fig. 2C, mean MDD: $p = 0.004$, $r^2 = 0.51$, Spearman’s correlation). This indicates that the MUA response magnitude (reflected in its mean response) has an impact on the decoding accuracy (see also, Astrand et al., 2015, Fig. 2).

### 2.9. Visuomotor functional properties of the recorded FEF neuronal population characterized by the memory-guided saccade task

In a majority of the recording sessions ($n = 13$ out of 15), both monkeys performed a memory-guided saccade task (for details see supplementary materials) directly prior to performing the cued detection task. In the memory-guided saccade task, visual stimuli were positioned at the same locations as in the cued detection task. By analyzing the MUA during this task, we categorized each channel as either pure visual (i.e. a significant MUA change 100 ms to 300 ms after the visual cue as compared to before the cue presentation), pure motor (i.e. a significant MUA change during saccade initiation, 100 ms to 300 ms after extinction of the fixation point), visuomotor (i.e. both visual and motor MUA changes) or none. A majority of the channels were categorized as visuomotor (49%) and only 12% and 17% as pure visual and pure motor channels, respectively (Supplementary Fig. S1). To relate the results of this work to the well-established functional properties of FEF neurons (Bruce and Goldberg, 1985), we investigated the correlation between channel categories (determined during the memory-guided saccade task) and spatial selectivity during the cued detection task. First, we found that the number of pure motor channels for each session does not impact decoding accuracy of target location during the cued detection task ($p = 0.45$), indicating that these channels do not carry any spatial information with regards to target location. Second, the number of pure visual and visuomotor channels correlates positively with both the number of spatially selective channels during the cued detection task ($p < 0.05$, $r^2 = 0.31$) and with decoding accuracy of target location during the cued detection task (S1; $p < 0.01$, $r^2 = 0.76$). These results strongly indicate that channels with visual properties during the memory-guided saccade task carries spatial information of the target in a cued target detection task.

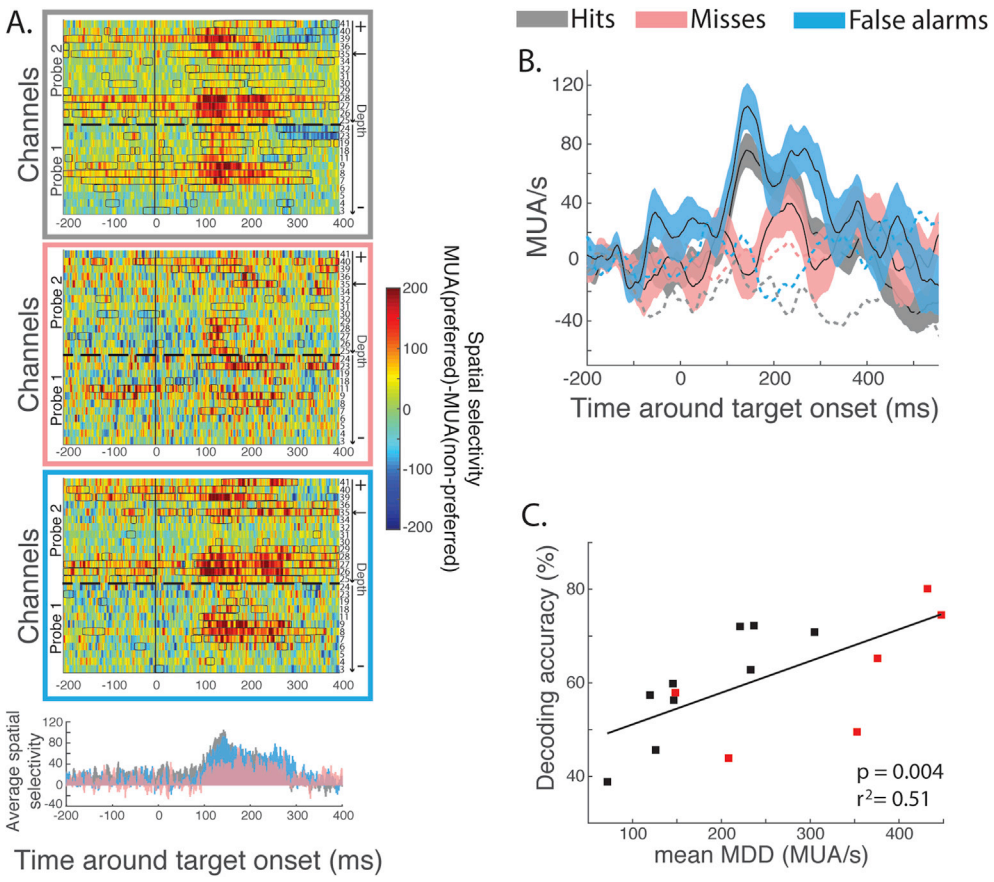



**Fig. 2. Multi-unit activity (MUA) spatial selectivity.** (A) Normalized average spatial selectivity (MUA/s; $MUA_{preferred\ location} - MUA_{non\text{-}preferred\ location}$) across trials for all channels with target-related activity (significant change in neuronal response 100–200 ms after target presentation as compared to −200 to −100 ms before target presentation, $p < 0.05$, Wilcoxon paired tests), for one session of monkey M2. Black boxes represent significant spatial selectivity, i.e. significantly different between responses at the preferred as compared to the non-preferred locations ($p < 0.05$, Wilcoxon tests). Hit (gray panel), miss (pink panel), and false alarm trials (blue panel) are shown for the same channels of the same session. The dashed horizontal lines indicate which probe the channels belong to and channel numbers are printed on the right side of each image along with a vertical arrow indicating the direction of depth of both probes. The horizontal arrow on the right side of each image indicate the channel that is plotted in B. PSTH for average spatial selectivity across channels are shown below, for each of the hit, miss and false alarm trials. (B) Average MUA per second (±s.e.) across trials for one channel during hit (gray), miss (pink), and false alarm trials (blue). MUA is plotted in time around target onset (for false alarms: in time around the distracter that evoked the response). Solid lines (resp. dotted) correspond to trials in which the target appeared in the preferred (resp. non-preferred) location. (C) For each session, the Euclidian multi-dimensional distance (MDD), computed in a time window of 100 ms to 200 ms after target presentation, is plotted against decoding accuracy (%) (black, monkey M1, red, monkey M2). The MDD is computed on mean MUA responses between all pairwise locations. The line represents the corresponding orthogonal regression for the dataset and Spearman's correlation statistics are indicated.

## 3. Results

### 3.1. Behavioral performance

Monkeys performed a spatially cued target detection task. Trials were divided according to their overt behavior as follows. A trial was considered correct (hit trial) if a response was produced between 200 ms and 700 ms from the onset of target luminance change. A trial in which the monkeys did not produce any response to a target was considered a miss. If a response was produced 200 ms to 700 ms following a distractor luminance change, this trial was considered a false alarm trial. M1 and M2 achieved 64.4% and 47.9% correct responses, respectively (Fig. 1C). Reaction times on these correct trials were on average 401 ms and 410 ms, respectively. Both monkeys produced more misses than false alarms (M1, Miss = 22.5%, FA = 12.9%; M2, Miss = 27.9%, FA = 23.0%, $p < 0.001$, Wilcoxon paired test) and reaction times were notably longer on false alarm trials compared to correct trials ($p < 0.01$ for both monkeys, Wilcoxon paired test).

### 3.2. Spatial selectivity of recorded multiunit activity

The MUA of each session was recorded on different days and hence in different locations within the FEF. Therefore, channel selectivity changed from one session to the next. Over all sessions, a majority of channels exhibited a significant target-related response, i.e. a significantly higher neuronal response 100–200 ms after target presentation as compared to pre-cue baseline (−200 to −100 ms pre-cue, Wilcoxon, $p < 0.05$, Monkey 1: 64% of channels; Monkey 2: 81% of channels). A large number of channels also exhibited a significant spatial selectivity, i.e. a significantly higher neuronal response for the preferred as compared to the non-preferred spatial position (Wilcoxon, $p < 0.05$), during target presentation (100–200 ms after target presentation, Monkey 1: 50% of channels; Monkey 2: 62% of channels). To provide a description of how the spatial selectivity of the recorded FEF neuronal population was distributed across the four target locations for each session we further calculated the number of spatially selective channels for each target location, and how many out of these were also spatially selective for the other target locations. Overall, the recorded FEF population on each session, had a balanced spatial representation of all target locations (for details see Supplementary Fig. S2).

Fig. 2A represents the spatial selectivity of individual target-related MUA channels (i.e. with a significant evoked response 100 ms to 200 ms after target presentation as compared to −200 ms to −100 ms before target presentation, Wilcoxon, $p < 0.05$) for one session of Monkey M2. The spatial selectivity was calculated for each channel as the difference between MUA, 100 ms to 200 ms after target presentation in the

preferred location (target location that elicited significant MUA maximal change from the pre-cue baseline) versus the non-preferred target location (target location that elicited minimal MUA change with respect to the pre-cue baseline). Spatial selectivity is shown in time from 200 ms before target presentation to 400 ms after target presentation. It is to be noted that target selective neurons can be identified along several millimeters in a row, confirming that recordings were being performed from the anterior bank of the arcuate sulcus. For each channel, time epochs of significant difference between the preferred and non-preferred channels are indicated by a black box. This information is represented for hit trials (Fig. 2A, top panel, gray outline), miss trials (Fig. 2A, middle panel, pink outline) and false alarm trials (Fig. 2A, bottom panel, blue outline, alignment is in this case on the time of presentation of the distractor that evoked the response). A substantial reduction in the evoked MUA after target onset can be observed during miss trials as compared to hit and false alarm trials (Fig. 2A, bottom histogram). Fig. 2B illustrates this for one specific MUA channel (arrow in Fig. 2A). Average MUA activity 100–200 ms after target presentation at the preferred location was significantly higher for hits (gray) and false alarms (blue) as compared to missed (pink) but was not significantly different between hits and false alarms. The neuronal activity was not modulated by target presentation at the non-preferred location (dashed lines). All subsequent decoding of single trial population spatial information is computed on this type of neuronal responses. The relationship between target position population decoding accuracy and single channel response statistics is described in the Methods (last section) and Fig. 2C.

### 3.3. Spatial representation of the cued location prior to target presentation (attention prioritization)

In the following, we quantify the available information about the cued location before target presentation in the prefrontal FEF neuronal ensembles being recorded from, as a function of the overt behavior of the monkeys. Specifically, recorded multi-unit activity (MUA), on correct trials, was used to train two regularized linear regressions (one along the x-axis and one along the y-axis) to associate bilateral neural response patterns at a given time from target presentation with the location of the expected target, and this for successive time windows around target presentation. Fig. 3A shows the resulting instantaneous classification performance, in time, around target presentation across all recording sessions. This classification performance represents the percentage of trials for which the classifier assigned its output to the cued quadrant. During correct trials (Fig. 3A, gray) before target presentation, classification performance is sustained at an average of 40%, well above the 95%-confidence limit. This is in accordance with (Astrand et al., 2016) in which we show that this increase in classification performance before target presentation represents attentional prioritization at the cued location. Accuracy before target presentation was sustained around 40% and was maintained above the 95% confidence interval limit (2-tailed non-parametric random permutation), indicating that the monkeys were actively orienting their attention to the cued spatial location. In contrast, for miss trials (Fig. 3A, pink) prior to target presentation, classification performance was around chance (25%) suggesting that, in these trials, the monkeys did not consistently orient their attention to the cued location, thus missing the target luminance change.

### 3.4. Spatial representation of the target

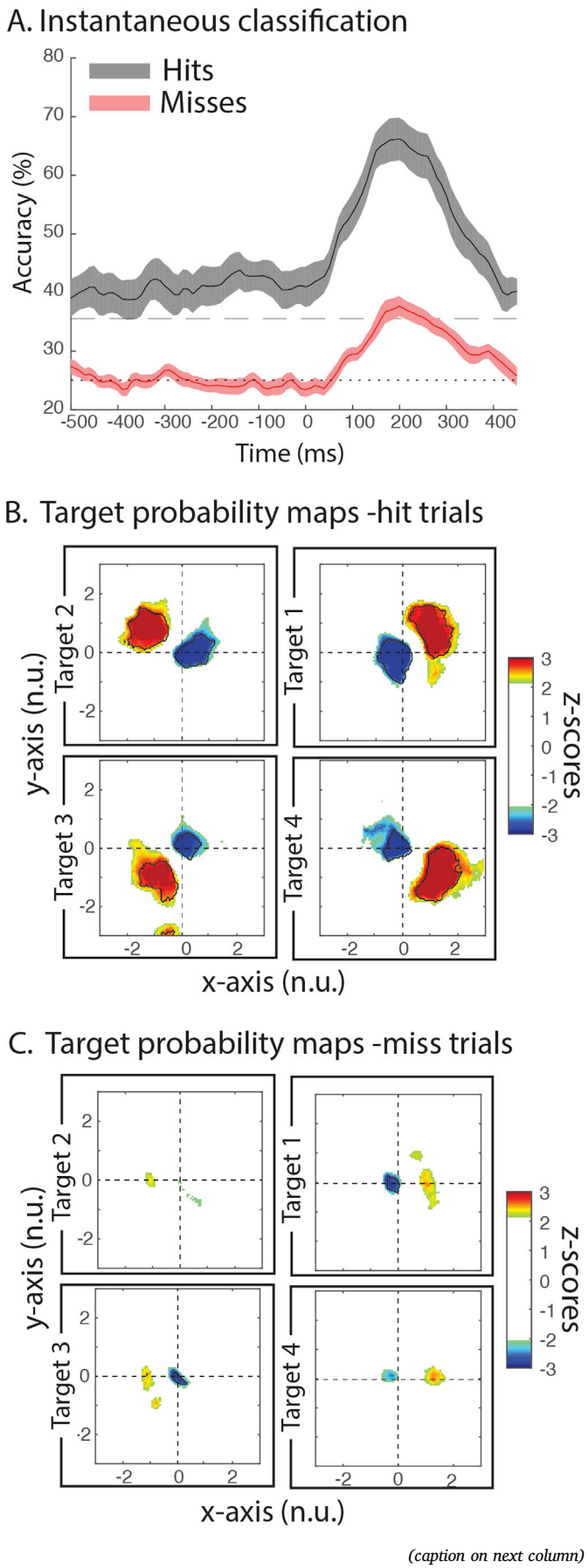
Instantaneous classification performance after target presentation (Fig. 3A) reflects how much information is available in the FEF population relative to target location. On correct trials, target-related information peaks at 200 ms, with an average of 65% correct predictions. This increase is well beyond the 95% confidence interval, and statistically significantly higher than the pre-target attentional orientation related signals (comparing accuracy 100 ms to 200 ms post-target onset, with accuracy 200 ms to 100 ms pre-target onset, average difference, 28%, $p < 0.001$, Wilcoxon paired test). On miss trials, significant increase across sessions is also observed although it barely reaches the upper 95% confidence interval limit (average difference 9%, p=<0.001, Wilcoxon paired test). This is in agreement with the findings that attentional orientation towards the receptive field of FEF neurons facilitates the detection or the discrimination of a target, showing a lower spiking rate on miss trials than on correct trials, both in the attention orientation period and following target presentation (Thompson and Schall, 1999; Ibos et al., 2013).

Whenever an item is perceived in our environment, it is implicitly associated to a more or less precise location in space. To investigate this aspect of target perception, the continuous output from the classifier was used to predict the spatial location of the reported target stimulus (target or distractor) using the same procedure that we have previously reported to track the attentional locus from FEF neuronal activities (Astrand et al., 2016). In the following, we focus on a time-period running from 100 ms to 200 ms after target presentation. Fig. 3B and 3C represent the (x,y) locations decoded from the population activities that were predicted statistically more (i.e. over-represented locations; red color scale, z-scores $> 2.33$, $p < 0.01$) or less frequently (under-represented locations; blue color scale, z-scores $< -2.33$, $p < 0.01$) as a function of the position of the target with respect to chance (estimated by a 2-tailed non-parametric random permutation test). On correct trials (Fig. 3B), a rather large area of overrepresented decoded locations around the location at which the target was actually presented can be observed for all four target locations (hot colors, Fig. 3B, average over all sessions, contour shows conjunction between all sessions and both monkeys). In contrast, a smaller area of underrepresented decoded locations can be observed around the fixation point for each target location (cold colors, Fig. 3B), suggesting that these locations were rarely spatially encoded by the monkeys. This is expected as neither targets nor distractors were ever presented at this location.

On miss trials (Fig. 3C, decoder trained on hit trials and tested on miss trials), only weakly overrepresented decoded locations can be identified. These locations fall short off target location and there is no conjunction between all sessions and the two monkeys. Underrepresented decoded locations can be identified around the fixation location, though they are smaller than those identified on correct trials (Fig. 3B). Overall, this indicates that correct target detection correlates with a reliable decoder localization of the target in the vicinity of its actual location. In contrast, missed targets correlate with an unreliable localization of the target. Overt perception, as assessed by behavior, and precise spatial information thus seem to be tightly related. To confirm this, we further probe this relationship between overt behavior and spatial information of distractor location on false alarm trials.

### 3.5. Spatial representation prior to distractor presentation

To investigate how the distractors, presented at a random time during the delay between cue and target presentation, were perceived in the different types of trials, the regularized linear regression was trained to associate MUA with the location of the target and tested onto distractor localization. Fig. 4A shows the instantaneous classification performance in time around distractor onset. On false alarm trials (Fig. 4A, blue), spatial information about the upcoming position of the distractor is significantly above the upper 95% confidence interval limit (two-tailed non-parametric random permutation). Because distractor location was completely pseudo-randomized across trials, there was no way for the monkeys to predict neither distractor probability of presentation, nor distractor time of presentation, nor distractor location. Thus, this spatial bias rather indicates that false alarm trials arise from misorientation of attention at the location before and at the time of presentation of the target (see Astrand et al., 2016 for a detailed analysis of this aspect). On correct (Fig. 4A, gray) and miss trials (Fig. 4A, pink), classification performance is well below chance (25%) almost reaching the lower 95% confidence interval limit before distractor presentation. At distracter




(caption on next column)

**Fig. 3. Target perception.** (A) Instantaneous classification accuracy of target location in time, aligned on target presentation (time = 0 ms). Graphs show mean classification accuracy (%) with associated standard error across sessions (n = 15) for hits (gray) and miss trials (red). At each time point the train-time of the classifier is the same as the test-time. The dotted horizontal line corresponds to chance classification performance (25%) and the striped horizontal line corresponds to the upper 95%-confidence interval determined by random permutation tests. (B–C) Decoded target (x,y)-location probability maps for hit trials (B) and miss trials (C). Maps represent z-scores of the spatial locations that are statistically over-represented (red color-scale, $p < 0.01$) or under-represented (blue color-scale, $p < 0.01$) with respect to the 95%-confidence interval (determined by random permutation tests) when decoding target location 100–200 ms after target presentation, over all sessions and both monkeys. Locations (10° or 13° eccentricity depending on the session) are normalized across sessions so that target locations are mapped to an eccentricity of ± 1. Each map corresponds to one of the four cued locations (e.g. upper right quadrant representing the map of spatial perception distribution for target appearing in the upper right quadrant). The black contour indicates the conjunction of mean z-scores over all sessions and both monkeys (i.e. common over- (under-, respectively) represented spatial locations across all sessions and both monkeys).

presentation during correct trials, the monkeys are correctly rejecting the distracter. Fig. 4A indicates that these correct rejections are behaving similar to miss trials, although a higher distractibility can be observed for miss trials (see Astrand et al., 2016 for a detailed analysis of this aspect). This indicates that, on these correct rejection and miss trials, attention is oriented away from the location of the upcoming distractor –thus accounting for it not being perceived on these trials. Thus, overall, these results suggest that the coincidence between distractor location and pre-distractor spatial attentional priority information contribute to overt behavior. As described below, this is mediated by an enhanced spatial representation of the distractor when it is selected by attention.

### 3.6. Spatial representation of distractor

Following distractor presentation, classification performance increases substantially during false alarm trials (Fig. 4A, blue) and reaches the same accuracy as for correct trials following target onset (correct trials, average accuracy from 100 to 200 ms post target onset: 61%, false alarm trials: 60%, $p = 0.64$, Wilcoxon paired test). On correct (gray) and miss (pink) trials, there is a substantially smaller increase in decoding accuracy following distractor presentation which remains below the upper 95% confidence interval limit (correct trials: average difference 7% below the 95% c.i., $p=<0.001$, miss trials: 11% below the 95% c.i., $p < 0.001$). In other words, as seen for reported target trials, when the distractor is reported (defined as overt false alarm behavior), a high decoding accuracy of distractor quadrant location in visual space can be observed. In contrast, for non-reported distractors (hits and misses) decoding accuracy is low.

To investigate how the distractor is spatially represented on the (x,y)-plane in the neuronal population, the continuous decoder output is analyzed as a function of overt behavior. Fig. 4B–D represent the (x,y) locations decoded from the population activities that were statistically overrepresented (red color scale, z-scores > 2.33, $p < 0.01$) or under-represented (blue color scale, z-scores < −2.33, $p < 0.01$) as a function of the position of the distractor with respect to chance (estimated by a 2-tailed non-parametric random permutation test). During both correct (Fig. 4B) and miss (Fig. 4C) trials, distractor location is either weakly or erroneously represented. In contrast, on false alarm trials (Fig. 4D), i.e. trials in which the distractor was mistakenly reported as a target, distractor location is significantly decoded at the actual location of the distractor, very much like what is shown for target localization on correct trials in Fig. 3B. Overall, this indicates that stimulus perception (whether a target or a distractor), assessed by overt behavior, correlates with a reliable stimulus localization that resides in the proximity of the real stimulus position.

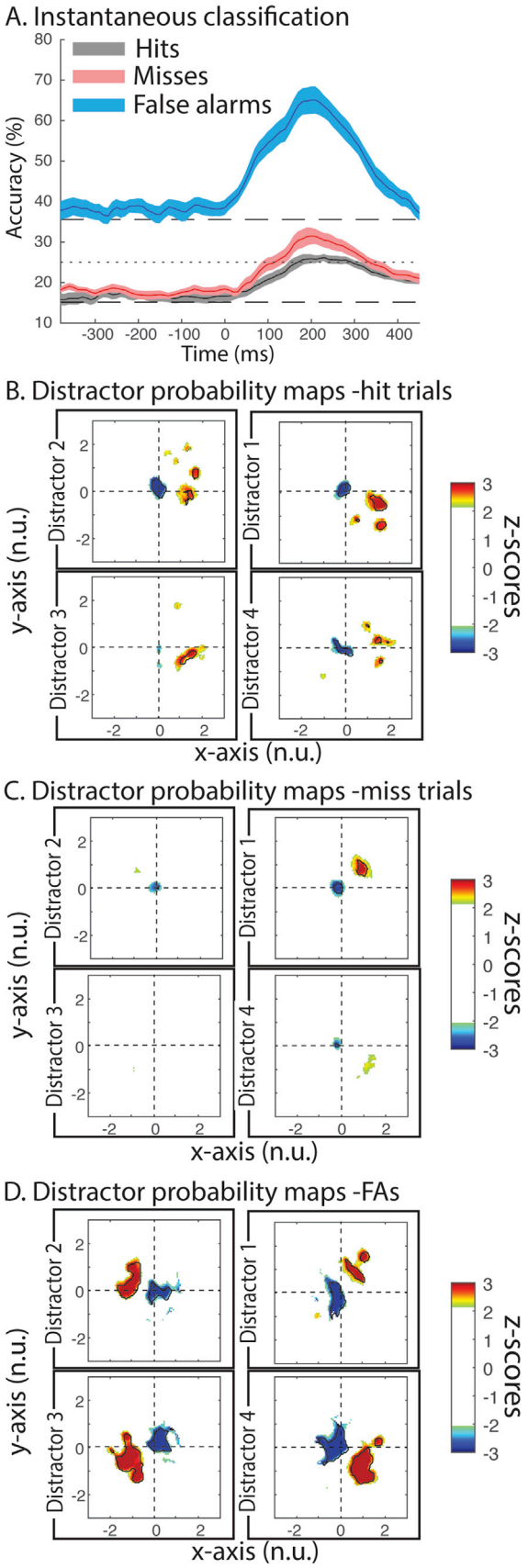



**Fig. 4. Distractor perception.** (A) Instantaneous classification accuracy of distractor location in time, aligned on distractor presentation (time = 0 ms). (B–D) Distractor (x,y)-location probability maps for hit trials (B), miss trials (C) and false alarm trials (D). All as in Fig. 3.

### 3.7. Distractor vs. target prefrontal representations

By task design, distractors and targets correspond to the same visual stimulus. What distinguishes one from the other is whether the spatial location where the stimulus is presented has been cued or not. In the following, we quantify the similarity between how the FEF encodes, at the population level, target and distractor spatial information. To this goal, a regularized linear regression was trained on correct trials, to discriminate between target and distractor, irrespective of stimulus position. Training was performed on MUA neuronal responses 100 ms to 200 ms following target and distractor presentation and decoding accuracy corresponded to how well the linear regression could predict that the presented stimulus was a target or a distractor, based on the observed neuronal activities. Fig. 5A shows decoding accuracies on novel correct (black), miss (pink) and false alarm (blue) test trials. Specifically, on correct trials, the classifier succeeded in discriminating the target from the distractor in 71% of instances (chance at 50%). On false alarm trials, decoding accuracies were in the same range as those observed on correct trials, and significantly above chance (69%, $p = 0.12$, Wilcoxon paired test, note that the distractor that evoked the response was considered to be the target in these trials). In contrast, on miss trials, decoding accuracy was significantly lower and hardly above chance (54%, $p < 0.001$, Wilcoxon paired test). Overall, this thus suggests that while information about the selected item is well represented in the FEF (target on hit trials

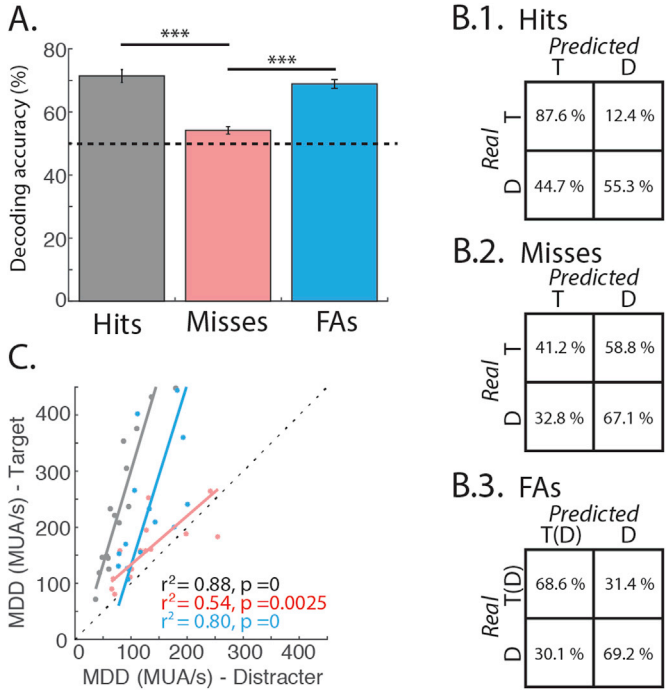



**Fig. 5. Target and distractor related information.** (A) Decoding accuracy at classifying target vs. distractor (mean+/-s.e., in %). Classifier is trained on hit trials and tested on hits (gray), misses (pink) and false alarm trials (blue). Accuracies are calculated over a 100–200 ms post-target or post-distractor time interval. Dotted horizontal line corresponds to chance level (50%). (B) Confusion matrices of the classification in A for hit trials (B.1), misses (B.2), and false alarms (B.3). Rows correspond to actual presented stimulus (distractor or target) and columns correspond to the predicted stimulus by the classifier. For false alarms, the target is taken as distractor that elicited the response T(D), all other stimuli are considered as distractors. (C) Euclidian multidimensional distance of MUA over all channels: target related response (y-axis) vs. distractor related response (x-axis) for hit (gray), miss (pink), and false alarm (blue) trials. Each dot corresponds to one session. Solid lines correspond to an orthogonal regression fit. Corresponding Spearman's correlation statistics are indicated.

and distractor on false alarm trials), hardly any information is available about the unselected item (target on miss trials).

The analysis of the confusion matrices refines this view (Fig. 5B). On hit trials (Fig. 5B1), while the target is correctly decoded as a target in 88% of instances, distractors are correctly decoded as distractors in only 55% of instances (88% vs 55%, $p < 0.001$, Wilcoxon paired test). In other words, the distractor is mistaken for a target in 45% of instances. On miss trials (Fig. 5B2), this pattern is reversed. Namely, the target is correctly detected as a target in only 41% of the trials (to be compared to the 88% correct classification on hit trials), while distractors are correctly decoded as distractors in up to 67% of instances. On false alarm trials (Fig. 5B3), both the decoded distractor (named T(D), due to the fact that it is selected as a target by the monkey) and the ignored distractor (D) are both correctly classified as target (69%) and distractor (69%) respectively. This suggests that, coexisting with the stronger selection process on hit and false alarm trials, described in the previous paragraph, an overall weaker distractor filtering process is at play during miss and false alarm trials, as available information about distractors is much higher on these trials than on hit trials. This possibly relates to higher noise correlations observed on miss and false alarm trials relative to hit trials that we previously described in the FEF neuronal population (Astrand et al., 2016).

The above describes differences in classification accuracy in target and distractor localization as a function of overt behavior and reflect complex changes in the joint neuronal population response patterns at stimulus presentation time. In order to gain a better understanding of how the spatial representations of a given visual stimulus, as assessed from the neuronal population response patterns, vary as a function of overt behavior, we proceeded as follows. We computed the multidimensional Euclidian distance in the multidimensional neuronal space (hereon simply called multidimensional distance, MDD) between the neuronal response patterns of the recorded populations (MUA) to the four possible stimuli locations for two conditions: after target and distractor presentation (see methods section). The MDD reflects how distant the MUA population response patterns is between the four stimuli positions. We first computed the MDD-target (Fig. 5C, y-axis), when the presented stimulus was an expected target (hits and misses) or a distractor that elicited a response (false alarms). We then computed the MDD-distractor (Fig. 5C, x-axis), when the presented stimulus was a non-reported distractor stimulus. For both computations, MDD was computed using neuronal activities averaged over 100–200 ms post-stimulus presentation. MDD-target and MDD-distractor are strongly correlated for all three types of trials (Fig. 5C, hits, gray: $p < 0.001$, $r^2 = 0.88$; misses, pink: $p < 0.001$, $r^2 = 0.80$; false alarms, blue: $p < 0.01$, $r^2 = 0.54$). This indicates that while neuronal responses varied when encoding target or distractor, these two processes were yet closely related. Furthermore, MDD is significantly higher following target presentation as compared to after distractor presentation for correct trials (average 238 MUA/s vs. 81 MUA/s, $p < 0.001$, Wilcoxon paired test). The same relationship can be observed for false alarm trials when considering the distractor that evoked the erroneous response as the target (average 244 MUA/s vs. 125 MUA/s, $p < 0.001$, Wilcoxon paired test). This indicates that stimuli evoking a behavior response result in more distant and dissociable response patterns to the four possible stimuli locations in the higher neuronal dimensional space. During miss trials there is a slight, although significant increase in the MDD following target presentation as compared to after distractor presentation (157 MUA/s vs. 127 MUA/s, $p < 0.01$). This may indicate the co-existence of two processes: 1) An active selection process associated with an overt behavioral response (in hits and false alarm trials) which coincides with enhanced neuronal responses following the visual stimulus, and 2) an active suppression mechanism that is associated with the absence of overt behavioral response and which coincides with decreased neuronal responses following a visual stimulus (i.e. compare neuronal responses following a distractor in hit trials with those following a target in miss trials).

In a last step, we sought to test whether the neuronal responses to the target and distractor, respectively, are associated with a change in the coefficient of variation of the neuronal responses. We thus calculated for each channel, on each session, the ratio between the standard deviation and the mean of MUA, following target or distractor presentation, for each trial type (hits, misses and false alarms). A 2-way ANOVA (trial type x target vs. distractor) revealed a significant main effect of target vs. distractor presentation ($p < 0.001$). Further post-hoc tests show that this coefficient of variation was significantly lower following target as compared to after distractor presentation for correct trials (target: 0.30, distractor: 0.32, $p < 0.001$, Wilcoxon paired test). This was also true during false alarm trials (distractor eliciting the response: 0.30, other distractors: 0.31, $p < 0.05$). During miss trials, no difference was observed (target: 0.31, distractor: 0.32, $p = 0.42$). In other words, stimulus selection was associated with more reliable neuronal responses and a 6% decrease in the coefficient of variation of the neuronal responses.

### 3.8. Behavioral correlates of target/distractor spatial representation

In Astrand et al. (2016), we show that the precision of the spatial representation of the cued location (i.e. same as target location), before target presentation, by the FEF neuronal population account for overt behavior: 1) the shorter the distance is between decoded location and the upcoming target, the lower the miss rates and the faster the reaction times; 2) the shorter the distance of the decoded location from the upcoming distractor, the higher the false alarm rates and the faster the reaction times. We argue that these correlations between the decoded location prior to target presentation and overt behavior strongly indicate that the decoded location reflects attentional orientation (Astrand et al., 2016). In this section, we analyze the correlation between the decoded location after target presentation and overt behavior and next we analyze how the decoded locations before and after target presentation relate to one another to optimize behavior. The distance between the decoded location after target presentation and the actual target location correlates with overt behavior in a similar manner as described for attention allocation (Fig. 6). Specifically, both monkeys show a significant correlation between median reaction times and mean distances between actual target location and (x,y) decoded target location in 20 equally sized distance bins (estimated on average neuronal responses from 100 to 200 ms after target onset), on correct trials (Fig. 6A, M1: $r^2 = 0.42$, $p < 0.001$, M2: $r^2 = 0.35$, $p < 0.001$). It is to be noted that in the task, the monkeys performed a manual response with the same hand throughout all sessions. As the response was identical for the four positions, we would not expect spatial information in the response preparation activity. Additionally, previous findings of single-unit activity in the FEF during a search task with a manual lever turn response, showed no activity related to the response indicating that FEF does not contribute to the manual behavioral report (Trageser et al., 2008). A significant correlation is also observed on false alarm trials between median reaction times and mean distances between actual distractor location (the distractor that evoked the response) and (x,y) distractor decoded location, in 20 equally sized distance bins, estimated on average neuronal responses from 100 to 200 ms after distractor onset (Fig. 6A, M1: $r^2 = 0.16$, $p < 0.05$, M2: $r^2 = 0.40$, $p < 0.001$). More specifically, for normalized distances below 1 and above 1, average reaction time substantially improved by 22 ms and 27 ms for monkey 1 and monkey 2, respectively. In addition, the proportion of misses over correct trials increases as the distance between target location and (x,y) target decoded location increases (Fig. 6B, M1: $r^2 = 0.79$, $p < 0.001$, M2: $r^2 = 0.79$, $p < 0.001$). In accordance, as the distance between distractor location (the distractor that evoked the response) and (x,y) distractor decoded location decreases, the proportion of false alarms over correct trials increases (Fig. 6B, M1: $r^2 = 0.77$, $p < 0.001$, M2: $r^2 = 0.93$, $p < 0.001$). These results establish an undisputable correlation between the decoded location of target and overt behavior of both monkeys. In other words, the spatial precision of which the target and distractors are represented in the FEF neuronal population strongly

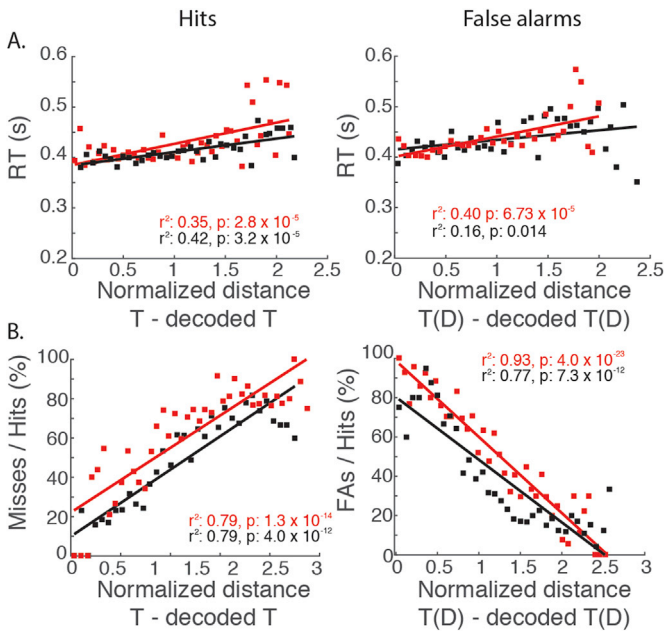



**Fig. 6. Reaction times (A) and detection performance (B) as a function of target to decoded target distance.** Target to decoded target distance are calculated on a time interval running from 100 to 200 ms post target presentation. For false alarms, the target is taken as the distractor that elicited the response T(D). Data are represented for hits (left panels), and false alarms (right panels). Each dot corresponds to the mean distance and median reaction times (A) or mean trial-type proportion rate (B) in each out of 20 equally sized distance bins (black, monkey M1, red, monkey M2). Data was fitted with an orthogonal regression (solid lines) and the corresponding statics of Spearman's correlation are indicated. Locations (10° or 13° eccentricity depending on the session) are normalized across sessions so that target locations are mapped to an eccentricity of ± 1.

influences both speed and accuracy of overt behavior.

### *3.9. Spatial representations before and after stimulus presentation – interactions towards optimal behavior*

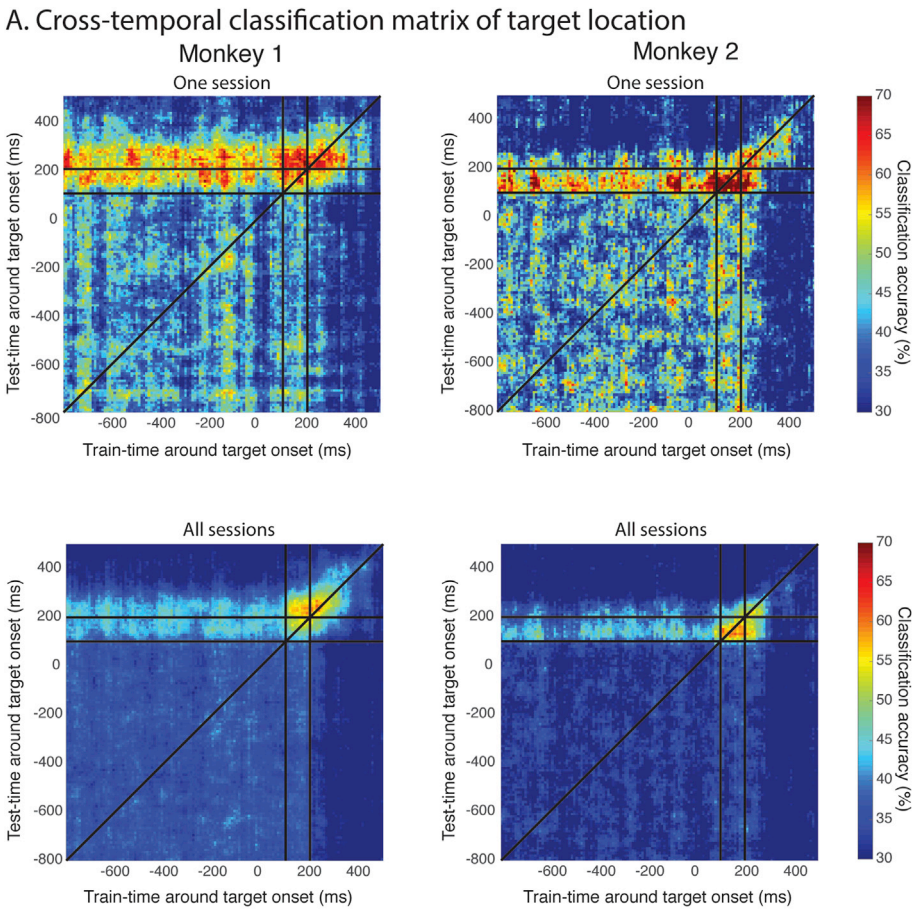
The spatial estimate of target location (as decoded from FEF neuronal activity 100 ms to 200 ms after target onset) is related to behavior in a very similar way as previously described for the spatial estimate of the cued location (i.e. attention orientation; 200 ms to 100 ms before target onset, Astrand et al., 2016). In a subsequent analysis, we examine the relation between attentional locus and spatial location of perception as estimated from the decoding of prefrontal neuronal population. We observe a strong correlation between the decoded location of attention and that of perception. More specifically, in correct trials during which attention was oriented close to the cued location prior to target presentation (attention located within a normalized distance of 0.7 from the cued location 300 ms to 200 ms prior to target presentation, hereafter referred to as close attention trials), the decoded location of the target (100 ms to 200 ms after target presentation) was significantly closer to its real physical location compared to trials during which attention was oriented outside the quadrant of the cued location (hereafter referred to as far attention trials) prior to target presentation (in close attention trials: the normalized distance between perceived location of target and its real physical location = 0.88±0.22, far attention trials: distance = 1.19±0.19, $p = 0.00036$). This indicates a strong relationship between pre-orientation of attention and the location of perception. In other words, the precision with which attention is pre-oriented to the cued location strongly influences how precise and veridical perceived target spatial location is.

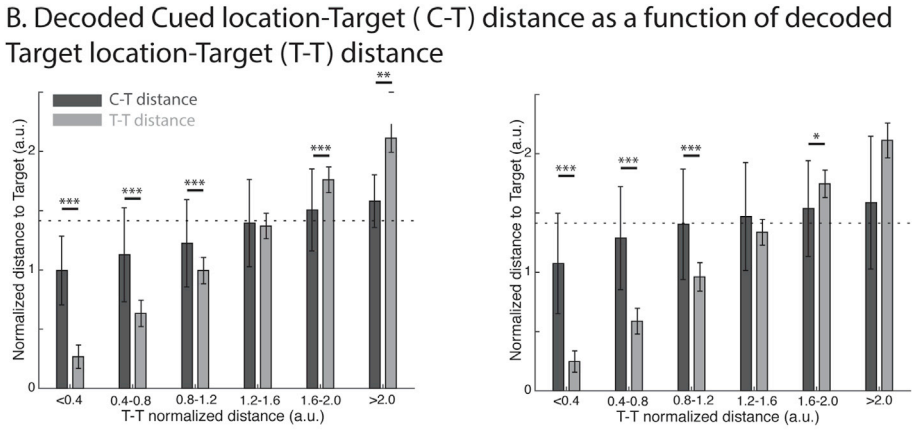
An important question is whether these estimates actually correspond the same neuronal process? If the spatial estimate of target location merely reflects a reorientation or reinforcement of spatial attention, we should expect a structurally similar neuronal pattern in the FEF population before and after target presentation, albeit a modulation in MUA amplitudes and/or reliability.

To shed light on this question, we performed a cross-temporal classification analysis of correct trials, using the discrete classification procedure, aligned on target presentation, and thus encompassing both before and after target-related processes (Fig. 7A, for details on classification procedure see sections related to Figs. 3A and 4A). In this analysis, data sets for training and testing the classifier were temporally dissociated to study whether different time points in the task relied on the same neural code (data were averaged in windows of 50 ms and slided with a step of 10 ms). From this analysis, two main observations can be highlighted: 1) the classification accuracy of target location using a classifier trained on data before target onset (i.e. attention epoch) increases substantially when tested after target location, and 2) the classification accuracy using a classifier trained on data after target onset decreases substantially when tested before target onset. These observations indicate that the attentional neural encoding before target onset is boosted by the target appearance, which may be interpreted as an attentional recapture. On top of this process, additional spatial information is activated by target appearance, the neuronal pattern of which is not shared with that of the encoding of the cued location prior to target presentation. We hypothesize that this additional spatial information specifically captures the spatial estimate of perception, independently of other processes in the task.

To further our understanding of how the spatial information represented in the neuronal population before and after target presentation are related to each other on a trial-by-trial basis, we investigated the decoded Cued location-Target (C-T) distance before target onset (estimated 200 ms to 100 ms before target onset) as a function of decoded Target location-Target (T-T) distance (estimated 100 ms to 200 ms after target onset, Fig. 7B). For both monkeys we show that while the C-T distance is shorter than $\sqrt{2}$ (dotted horizontal line; when C-T$>\sqrt{2}$ the decoded cued location is within a non-cued quadrant of the screen), the T-T distance is significantly shorter than the C-T distance (M1: $p < 0.001$, M2: $p < 0.001$, Wilcoxon signed rank test with Bonferroni correction of $n = 6$ tests). In contrast, for trials in which the decoded cued location is outside the cued quadrant (C-T distance$>\sqrt{2}$), the T-T distance is significantly longer than C-T distance (M1: $p < 0.01$, M2: $p < 0.01$ for $1.6 < PT < 2.0$ and $p = 0.56$ for $PT > 2.0$, Wilcoxon signed rank test with Bonferroni correction of $n = 6$ tests). As these results strengthen our hypothesis that different neural processes underlie the decoded locations before and after target presentation, they also show that spatial information of an upcoming target (i.e. spatial attention), encoded in the neuronal population, strongly impacts the accuracy of the spatial representation of the target.

Relating the C-T- and T-T-distances to response times (Fig. 7C) reveals that the spatial information in the neuronal population before and after target presentation will impact response times differently for the two monkeys. Specifically, linearly fitting the 3-dimensional data points (C-T-distance x T-T-distance x Response time) for both monkeys showed that for monkey 1, the T-T distance better explained the response times ($weight_{Attention} = 2.23$, $weight_{Perception} = 28.99$), and for monkey 2, both the C-T and T-T distances explained the response times ($weight_{Attention} = 19.61$, $weight_{Perception} = 33.83$). These results indicate that the decoded target location (after target presentation) reflects a spatial estimate of perception, differently encoded than the spatial estimate of attention prior to target presentation. These observations thus indicate a tight relationship between the spatial encoding of attention and perception, however, the exact relationship might depend on internal strategies rather than on a strict causal relationship.



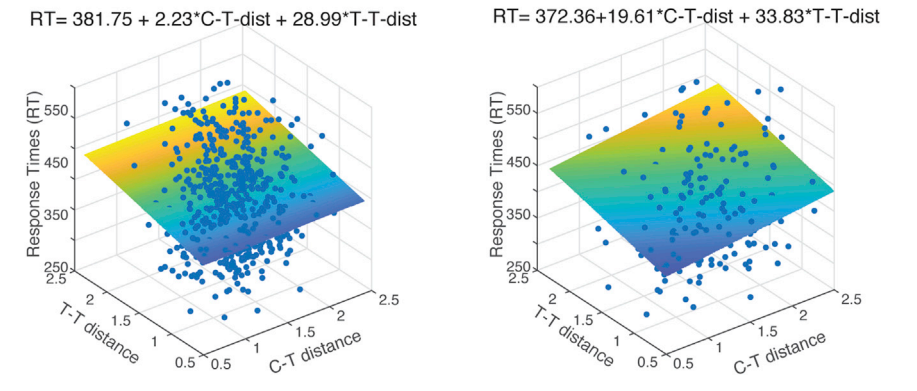




**Fig. 7. Spatial representation before and after target presentation.** (A) Cross-temporal classification matrices of target location for monkey 1 (left column) and monkey 2 (right column) and for one example session (upper matrices) and across all sessions (lower matrices). Colors represent classification accuracy (with red as the highest accuracy) as a function of classifier train-time during the task (x-axis) and classifier test-time (y-axis). Black lines correspond to the diagonal and the defined perception epoch (100 ms to 200 ms after target onset). (B) Decoded Cued location-Target distance (C-T, estimated 200 ms to 100 ms before target onset) as a function of decoded Target location-Target distance (T-T, estimated 100 ms to 200 ms after target onset) for monkey 1 (left) and monkey 2 (right). The normalized distance between the spatial estimate of the target and the physical local of the target (T-T distance) are represented in 6 bins (light gray bars) with the corresponding distance between the spatial estimate of the cued location and that of the target (C-T distance, dark gray bars). The dotted horizontal line corresponds to a distance of $\sqrt{2}$. Asterisks denote statistical difference between distributions (***$p < 0.001$, **$p < 0.01$, and *$p < 0.05$). (C) Linear modeling of reaction times as a function of the optimal weighted sum of C-T and T-T distances for monkey 1 (left) and monkey 2 (right). For each trial, the C-T and T-T distances are plotted against the response time. Data was linearly fitted and the resulting hyperplane is visualized in each 3-D plot with yellow colors corresponding to slower response times.

### 3.10. Population target location estimates reflect onto underlying MUA

The distance between the actual visual stimuli and where they are represented in space as decoded from the FEF neuronal responses is a population estimate. Here, we probe how much this population estimate reflects onto the underlying MUA responses. In other words, do MUA responses correlate with decoder output? Fig. 8 represents the correlation between the normalized amplitude of MUA and the distance between target (distractor evoking a response for false alarm trials) location and (x,y) decoder output. MUA signals that had a significant positive modulation of their response, on correct trials, in the 100–200 ms time window following target presentation, as compared to the −200 to −100 ms prior to target presentation (Wilcoxon paired tests, $p < 0.05$), were selected for further analysis (508 out of 720 MUA channels and 960 out of 2880 target-related responses). The target-to-decoder output distance was averaged within 20 equally sized distance bins and the MUA from the corresponding trials within each bin was averaged. On correct trials (Fig. 8, gray), a significant correlation between target-decoder output distance and MUA amplitude can be observed ($r^2 = 0.48$, $p < 0.01$). On false alarm trials, overall MUA amplitude is lower compared to hit trials ($p < 0.001$, Wilcoxon test) but a significant correlation can also be observed between MUA amplitude and distractor-decoder output distance ($r^2 = 0.32$, $p = 0.01$). On miss trials, the MUA amplitude is substantially lower as compared to hit trials (485 MUA/s vs. 360 MUA/s, $p < 0.001$, Wilcoxon test). For these trials, a trend towards significance can be observed between the correlation between MUA amplitude and target-decoder output distance ($r^2 = 0.23$, $p = 0.06$). Overall, FEF MUA activity thus reflects the decoded stimulus location, rather than actual stimulus physical location.

## 4. Discussion

Visual perception is defined as the conscious representation of a visual item. It is experimentally assessed by requesting an overt report by the subject that can take different forms: a detection, a discrimination, a verbal report etc. In this work, we identify the neuronal prefrontal correlates of visual perception from a neuronal population perspective (Astrand et al., 2014, 2015, 2016). The neurons in the FEF have traditionally been analyzed with univariate techniques, such as correlating single-unit recording activity to individual task aspects. These techniques have enabled to identify highly distinct roles of individual neurons in the FEF. Specifically, early studies have established the well-known categorization of FEF neurons into visual, motor or visuomotor neurons based on their activation pattern during the memory-guided saccade task (Bruce and Goldberg, 1985). However, these univariate analyses fail to explain the often-observed complex activation patterns in neurons in higher-order brain areas (FEF included). Through advances in multivariate pattern analysis, it has become possible to study neuronal population encoding (Buonomano and Maass, 2009), which has provided evidence for prefrontal neurons exhibiting selectivity for multiple relevant aspect of a given task (so-called a mixed selectivity, Fusi, 2016; Rigotti et al., 2013). Rigotti et al. (2013) specifically show that prefrontal neuronal populations encode distributed information not detectable using univariate analysis. These findings emphasize the difficulty in linking univariate functional properties of FEF neurons to the multivariate decoding analysis. Here, we propose that the neuronal population, with a major contribution of visual and visuo-motor neurons, encode a perceived position of visual stimuli that may be different from their veridical position. This encoded perceived position drives overt behavior, in partial interaction with spatial orientation signals and cannot be captured by individual neurons. This framework is novel and contrasts with, though also complementing previous studies on the contribution of FEF neuronal responses to perception. Specifically, we show that when a stimulus is reported, whether this stimulus is the target of behavior or a distractor, the neuronal FEF population precisely encodes its location. In contrast, when a stimulus is not reported, whether this stimulus is the target of behavior or a distractor, its location, as encoded by the FEF, does not match its actual location. We describe a strong correlation between the error in the estimation of the position of the visual stimulus in space as coded by the neuronal population with respect to its actual physical location, and overt behavior. We further show that the spatial estimate of the visual stimulus is partly independent from the spatial location of attention prior to the stimulus onset, and that each monkey relied differently upon the spatial encoding of attention and perception in their behavioral optimization. These observations are discussed below.

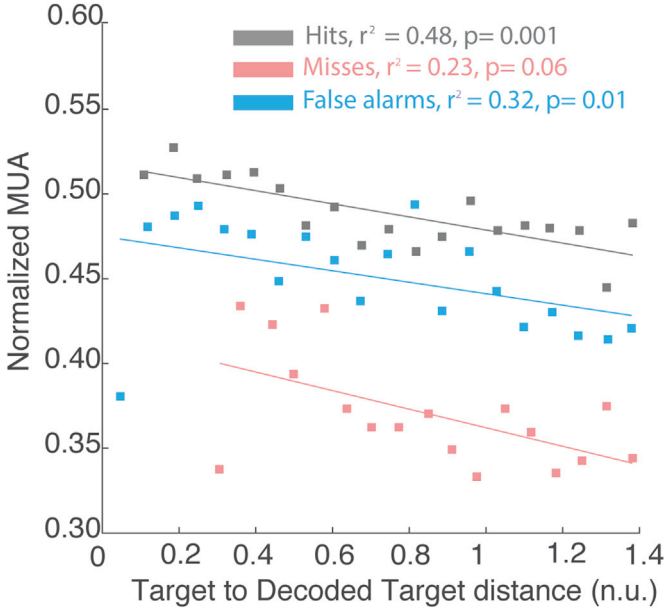



**Fig. 8. Normalized MUA as a function of target to decoded target distance.** Normalized MUA and target to decoded target distance are calculated on a time interval running from 100 to 200 ms post target presentation. Data cumulated for both monkeys are represented for hits (gray), misses (pink) and false alarms (blue). In the case of false alarms, measures are extracted relative to distractor presentation. All else as in Fig. 6.

### 4.1. Precise stimulus location estimation correlates with stimulus selection

In the absence of attentional pre-orientation (cued), single-unit neuronal activity in the FEF during pop-out visual search, driven by bottom-up mechanisms, has been shown to reliably encode the presence of a visual stimulus within the neuron's receptive field (RF) independently of whether a correct or erroneous manual report is produced (Trageser et al., 2008). This indicates that, in this task, the FEF does not contribute to the behavioral report. In contrast, on a difficult cued target detection task, driven by top-down mechanisms, the speed and accuracy of the behavioral response on individual trials is predicted by the magnitude of single neuron responses to the target when presented in the neuron's RF, this magnitude being lower both on error trials and on trials with longer response times (Monosov and Thompson, 2009). Single-unit analyses have provided great insight into how individual neurons behave and encode information but interactions between neurons within a population are lost (e.g. spatial attention (Cohen and Maunsell, 2009; Astrand et al., 2016)).

At the neuronal population level, we show very high classification rates of both targets and distractors when these are selected to produce a behavioral response (Figs. 3, 4, hits and false alarm trials). This translates into a confined localization on the decoding probability maps around the actual physical location of the target/distractor. In other words, on these trials, the location of the target or of the distractor is precisely

represented in the prefrontal cortex. In contrast, classification rates of both targets and distractors are below the 95% confidence interval when these are not selected for behavioral report (misses and correct rejections). This corresponds to an unreliable localization on the target/distractor decoding probability maps. In other words, visual stimulus selection correlates with a reliable neuronal population representation of stimulus location on a trial-to-trial basis. The decoding probability maps (Figs. 3, 4) also show an underrepresented area in the vicinity of the fixation point. This indicates that monkeys avoid selecting this portion of space as compared to other spatial locations, however no conclusions can yet be drawn as to whether active suppression mechanisms are at play.

### 4.2. Neuronal population information accounts for overt behavior

As discussed in the previous section, in cued target detection tasks, the speed and accuracy of the overt response on individual trials is predicted by the magnitude of single neuron responses to a target in the RF (Monosov and Thompson, 2009). Here we show that the spatial estimate of target/distractor location provided by the FEF neuronal population accounts for speed and accuracy of overt behavior. Specifically, an accurate spatial representation coincides with 1) shorter reaction times, 2) lower proportions of misses and 3) higher proportions of false alarms when the location of the target/distractor is estimated at its veridical position as compared to further away (Fig. 6). In other words, the population estimate of target/distractor location parametrically accounts for behavior, with an explained variance of miss and false alarm rates ranging between 70% and 90%. Multi-unit activity in response to the target/distractor presentation also co-varies with this population estimate. However, in this case, the explained variance is much lower and ranges between 20% (MUA response to target in misses) and 50% (MUA response to target in hits). This indicates that the neuronal population better accounts for overt behavior than single neuron or multi-unit activity.

### 4.3. Target selection vs. distractor filtering

Suzuki and Gottlieb (2013) show that the dorsolateral prefrontal cortex single-unit neural spiking activity following a distractor (that is identical to the target) is positively correlated to error rates. On a population level we corroborate and extend this finding by showing that behavioral performance, in terms of accuracy and response times, correlates with the spatial representation of a distractor in the FEF (Fig. 6). Specifically, we show that as the distance between the FEF neuronal population estimate of the locus of distractor and the actual position of the distractor decreases (i.e. as the error of the spatial estimation decreases), the false alarm rate increases and response times during false alarm trials decrease. We further observe that multiunit activity following distractor presentation negatively correlates with the error between the estimated distractor location in false alarm trials and its actual physical location (Fig. 8). This indicates that the selection of a distractor for behavioral report co-varies with this distance error measure. The shorter the distance the more likely the selection. Likewise, target selection co-varies in a similar manner with the error between the estimated target location in hit trials and its actual physical location. We similarly observe that the multiunit activity following target presentation negatively correlates with this distance error measure.

Hit trials correspond to trials in which the target has been selected. In contrast, false alarm trials correspond to trials in which a distractor failed to be filtered. By task design, target and distractors were identical physical stimuli, the only distinguishing factor being whether the stimulus was presented at the cued location or not. As a result, one expects that, when perceived, targets and distractors would be represented in identical manners. Average normalized multi-unit activity in response to a detected target or to a detected distractor were significantly higher than the average normalized multi-unit activity in response to a missed target. Average normalized multi-unit activity in response to a detected target was only slightly stronger than the average normalized multi-unit activity in response to a detected distractor. The neuronal population does not discriminate between an unselected target and an unselected distractor. Likewise, the neuronal population equally discriminates between selected and unselected target and selected and unselected distractors. In this context, it is worth noting that we show that target selection corresponds to a lower trial-to-trial MUA variability compared to distractor selection, suggesting a higher degree of uncertainty during distractor selection. Previous studies have shown that the trial-to-trial Fano factor (variance/mean) is similar for both targets and distracters appearing in the RF of FEF neurons (Bichot et al., 2001; Purcell et al., 2012). However, these studies only considered correct trials, i.e. trials on which targets were correctly detected and distractors correctly discarded. In contrast with what we describe, this FEF change in Fano factor does not reflect stimulus selection. Overall, our results suggest that once a stimulus has been selected, targets and distractors are to a great extent undistinguishable to the neuronal population though distractor selection process is more subject to variability. As shown in Astrand et al. (2016), the erroneous selection of a distracter is most probably due to a misorientation of attention during initial stages of the task (Heitz et al., 2010; Astrand et al., 2016).

### 4.4. Interactions between spatial attention and perception

During low signal or high noise conditions, spatial attention has been shown to facilitate perception at the locus of attention. Indeed, behavioral responses to attended stimuli are faster (Yantis and Jonides, 1990) and visual sensitivity at attended locations is enhanced (Bashinski and Bacharach, 1980; Carrasco, 2011). At the neuronal level, attention has been proposed to operate through a variety of mechanisms including enhanced neuronal response to visual stimuli when attention is oriented towards the receptive field of the neuron (e.g. McAdams and Maunsell, 1999); a shrinkage of visual receptive fields (RF) and a shift towards the attended location (Ben Hamed et al., 2002; Womelsdorf et al., 2006, 2008); a decreased trial-to-trial variability of individual neuron's response (Cohen and Maunsell, 2009); an increased synaptic efficacy (Briggs et al., 2013); a decrease in noise correlations between neurons in extrastriate visual areas (Cohen and Maunsell, 2009) as well as in prefrontal area FEF (Astrand et al., 2016, but see also Cohen et al., 2010), and decreased neuronal response latencies (Galashan et al., 2013). This is proposed to have as overall effect to enhance perceptual processing, possibly through local (Chalk et al., 2010; Panagiotaropoulos et al., 2012) and long-range (Popov et al., 2017) neuronal coupling mechanisms. Furthermore, perception has traditionally been described as an "on-off" state; a visual stimulus is either perceived or not perceived. The reason why a stimulus fails to be consciously perceived has mainly been explained by classical models (e.g. Signal Detection Theory (SDT) (Green and Swets, 1966)) assuming that the stimulus-evoked activity has to reach a certain threshold in order for it to be consciously perceived (Ratcliff and McKoon, 2008). Recent findings have extended this view to suggest that the stimulus-induced activity must propagate to higher-order prefrontal areas and reach a threshold sufficient to activate "ignition", an event causing self-sustained prefrontal neuronal activation to broadcast information between many brain areas (Vugt et al., 2018). Although a threshold might explain why weak near-threshold stimuli reach consciousness, it does not address whether the perceived location of the stimulus is related to overt behavior (i.e. perception contains a spatial component influencing behavior). Here, we use multi-variate pattern analyses from simultaneously recorded multi-unit activities, thus capturing information encoded in the FEF neuronal population, and we specifically show that overt behavior (speed and accuracy) can be explained by the spatial accuracy of the reported stimulus encoded by the neuronal population. We further show that the spatial encoding neuronal pattern of the reported stimulus is partly independent from that of the spatial attention orientation, indicating that perception contains a spatial component de-correlated from pre-existing attention signals.

The correlations we describe between overt behavior and prefrontal target-related spatial representations could be interpreted as either a reinforcement/recapture of spatial attention or change in the strength of the percept associated with the target or the distractor rather than as a change in the estimate of its spatial position; a strong percept at the time of target correlating with higher probability of correct detections and a strong percept at the time of distractor correlating with higher probability of false alarms. Two arguments speak against this. First, the classification we are applying is not discriminating perception vs. failed perception trials, but rather associating the observed neuronal activities to a spatial estimate. Second, and most importantly, an asymmetry in the cross-temporal classification analyses (Fig. 7A) strongly suggests that additional spatial information is present after target onset that is encoded differently in the neuronal population than spatial attention before target onset. We argue that the spatial information after target onset, being strongly correlated to overt behavior, is evidence for a spatial estimate of perception. While partly independent from spatial attention, we describe a close link between the spatial estimates of attention and perception (Fig. 7B and C). The closer attention is to the cued location prior to target onset, the more spatially accurate the neuronal population will encode the target. As long as attention is within the cued quadrant of the screen prior to target onset, the spatial estimate of perception will be significantly more accurate than that of attention. However, even though monkeys correctly detected the target, when attention is outside of the cued quadrant prior to target onset, the spatial estimate of perception will be encoded even farther away from the physical location of the target. Rather than a strict causal relationship between spatial attention and perception for optimizing behavior, we show that response times can be explained by a weighted balance between the spatial estimates before (attention) and after target onset (perception), of which weights may depend on individual strategy or task difficulty.

In conclusion, we show that the population neuronal responses in the FEF not only inform on whether a stimulus has been perceived or not, but also on how accurately it was localized in space, irrespectively of whether the stimulus was an actual target of behavior or an irrelevant stimulus. The accuracy of this spatial representation strongly correlates with overt behavior in terms of response time and accuracy. A strong prediction of this is that in a cued-target detection task in which a spatial response is required (e.g. saccade or pointing), overt error will correlate with the internal prefrontal representation of target location. From a fundamental perspective, while perception is often viewed as an all or nothing variable, our work provides evidence for a measure of reliability of the percept: when a stimulus is detected, this detection can be associated with a very good spatial estimate or with a poor spatial estimate. This view challenges classical models of decision-making or at least calls for the integration of this spatial dimension. From an applied perspective, understanding the neuronal population substrates of stimulus selection, distractor filtering and overt behavior is crucial for developing novel technological advances to improve abilities related to visual discrimination and selection of relevant information in noisy environments or in pathological conditions (Astrand et al., 2014).

## Author contribution statement

Conceptualization, S.B.H.; Methodology, E.A., and S.B.H.; Investigation, E.A., C.W.; Writing – Original Draft, S.B.H. and E.A.; Writing – Review & Editing, S.B.H. and E.A.; Funding Acquisition S.B.H. and E.A.; Supervision, S.B.H.

## Funding

This work was supported by the Centre National de la Recherche Scientifique (CNRS), Direction Générale de l'Armement (DGA), Fondation pour la Recherche Médicale (FRM), and French National Research Agency (ANR-11-BSV4-0011, ANR-11-LABX-0042, ANR-11-IDEX-0007).

## Acknowledgements

We thank research engineer Serge Pinède for technical support and Jean-Luc Charieau and Fabrice Hérant for animal care. All procedures were approved by the local animal care committee (C2EA42-13-02-0401-01) in compliance with the European Community Directive 2010/63/UE on Animal Care.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2020.116517.

## References

Astrand, E., Enel, P., Ibos, G., Dominey, P.F., Baraduc, P., Ben Hamed, S., 2014. Comparison of classifiers for decoding sensory and cognitive information from prefrontal neuronal populations. PLoS One 9.
Astrand, E., Ibos, G., Duhamel, J.-R., Ben Hamed, S., 2015. Differential dynamics of spatial attention, position, and color coding within the parietofrontal network. J. Neurosci. 35, 3174–3189.
Astrand, E., Wardak, C., Baraduc, P., Ben Hamed, S., 2016. Direct two-dimensional access to the spatial location of covert attention in macaque prefrontal cortex. Curr. Biol. 26 (13), 1699–1704.
Bashinski, H.S., Bacharach, V.R., 1980. Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. Percept. Psychophys. 28, 241–248.
Ben Hamed, S., Duhamel, J.-R., Bremmer, F., Graf, W., 2002. Visual receptive field modulation in the lateral intraparietal area during attentive fixation and free gaze. Cerebr. Cortex 12, 234–245.
Bichot, N.P., Thompson, K.G., Chenchal, R.S., Schall, J.D., 2001. Reliability of macaque frontal eye field neurons signaling saccade targets during visual search. J. Neurosci. 713–725.
Briggs, F., Mangun, G.R., Usrey, W.M., 2013. Attention enhances synaptic efficacy and the signal-to-noise ratio in neural circuits. Nature 499, 476–480.
Bruce, C.J., Goldberg, M.E., 1985. Primate frontal eye fields. I. Single neurons discharging before saccades. J. Neurophysiol. 53, 603–635.
Buonomano V., D., Maass, W., 2009. State-dependent computations: spatiotemporal processing in cortical networks. Nat. Rev. Neurosci. 10 (2), 113–125.
Carrasco, M., 2011. Visual attention: the past 25 years. Vis. Res. 51, 1484–1525.
Chalk, M., Herrero, J.L., Gieselmann, M.A., Delicato, L.S., Gotthardt, S., Thiele, A., 2010. Attention reduces stimulus-driven gamma frequency oscillations and spike field coherence in V1. Neuron 66, 114–125.
Cohen, J.Y., Crowder, E.A., Heitz, R.P., Subraveti, C.R., Thompson, K.G., Woodman, G.F., Schall, J.D., 2010. Cooperation and competition among frontal eye field neurons during visual target selection. J. Neurosci. 30, 3227–3238.
Cohen, M.R., Maunsell, J.H.R., 2009. Attention improves performance primarily by reducing interneuronal correlations. Nat. Neurosci. 12, 1594–1600.
Farbod Kia, S., Åstrand, E., Ibos, G., Ben Hamed, S., 2011. Readout of the intrinsic and extrinsic properties of a stimulus from un-experienced neuronal activities: towards cognitive neuroprostheses. J. Physiol. Paris 105, 115–122.
Fusi, S., et al., 2016. Why neurons mix: high dimensionality for higher cognition. Curr. Opin. Neurobiol. 37, 66–74.
Galashan, F.O., Saßen, H.C., Kreiter, A.K., Wegener, D., 2013. Monkey area MT latencies to speed changes depend on attention and correlate with behavioral reaction times. Neuron 78, 740–750.
Green M., D., Swets A., J., 1966. Signal detection theory and psychophysics. John Wiley, Oxford, England.
Heitz, R.P., Cohen, J.Y., Woodman, G.F., Schall, J.D., 2010. Neural correlates of correct and errant attentional selection revealed through N2pc and frontal eye field activity. J. Neurophysiol. 104, 2433–2441.
Ibos, G., Duhamel, J.-R., Ben Hamed, S., 2013. A functional hierarchy within the parietofrontal network in stimulus selection and attention control. J. Neurosci. 33, 8359–8369.
Libedinsky, C., Livingstone, M., 2011. Role of prefrontal cortex in conscious visual perception. J. Neurosci. 31, 64–69.
McAdams, C.J., Maunsell, J.H., 1999. Effects of attention on the reliability of individual neurons in monkey visual cortex. Neuron 23, 765–773.
Monosov, I.E., Thompson, K.G., 2009. Frontal eye field activity enhances object identification during covert visual search. J. Neurophysiol. 102, 3656–3672.
Monosov, I.E., Trageser, J.C., Thompson, K.G., 2008. Measurements of simultaneously recorded spiking activity and local field potentials suggest that spatial selection emerges in the frontal eye field. Neuron 57, 614–625.
Moore, T., Fallah, M., 2004. Microstimulation of the frontal eye field and its effects on covert spatial attention. J. Neurophysiol. 91, 152–162.
Panagiotaropoulos, T.I., Deco, G., Kapoor, V., Logothetis, N.K., 2012. Neuronal discharges and gamma oscillations explicitly reflect visual consciousness in the lateral prefrontal cortex. Neuron 74, 924–935.
Popov, T., Kastner, S., Jensen, O., 2017. FEF-controlled alpha delay activity precedes stimulus-induced gamma-band Activity in visual cortex. J. Neurosci. 37, 4117–4127.
Purcell, B.A., Heitz, R.P., Cohen, J.Y., Schall, J.D., 2012. Response variability of frontal eye field neurons modulates with sensory input and saccade preparation but not visual search salience. J. Neurophysiol. 108, 2737–2750.

Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 20, 873–922.
Rigotti, M., et al., 2013. The importance of mixed selectivity in complex cognitive tasks. Nature 497 (7451), 585–590.
Sato, T.R., Watanabe, K., Thompson, K.G., Schall, J.D., 2003. Effect of target-distractor similarity on FEF visual selection in the absence of the target. Exp. Brain Res. 151, 356–363.
Schall, J.D., Hanes, D.P., 1993. Neural basis of saccade target selection in frontal eye field during visual search. Nature 366, 467–469.
Suzuki, M., Gottlieb, J., 2013. Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. Nat. Neurosci. 16, 98–104.
Thompson, K.G., Biscoe, K.L., Sato, T.R., 2005. Neuronal basis of covert spatial attention in the frontal eye field. J. Neurosci. 25, 9479–9487.
Thompson, K.G., Hanes, D.P., Bichot, N.P., Schall, J.D., 1996. Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. J. Neurophysiol. 4040–4055.
Thompson, K.G., Schall, J.D., 1999. The detection of visual signals by macaque frontal eye field during masking. Nat. Neurosci. 2, 283–288.
Thompson, K.G., Schall, J.D., 2000. Antecedents and correlates of visual detection and awareness in macaque prefrontal cortex. Vis. Res. 40, 1523–1538.
Trageser, J.C., Monosov, I.E., Zhou, Y., Thompson, K.G., 2008. A perceptual representation in the frontal eye field during covert visual search that is more reliable than the behavioral report. Eur. J. Neurosci. 28, 2542–2549.
Tremblay, S., Pieper, F., Sachs, A., Martinez-Trujillo, J., 2015. Attentional filtering of visual information by neuronal ensembles in the primate lateral prefrontal cortex. Neuron 85, 202–215.
Vugt, B van, Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., Roelfsema, P.R., 2018. The threshold for conscious report: signal loss and response bias in visual and frontal cortex. Science 360, 537–542.
Wardak, C., Ibos, G., Duhamel, J.-R., Olivier, E., 2006. Contribution of the monkey frontal eye field to covert visual attention. J. Neurosci. 26, 4228–4235.
Womelsdorf, T., Anton-Erxleben, K., Pieper, F., Treue, S., 2006. Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. Nat. Neurosci. 9, 1156–1160.
Womelsdorf, T., Anton-Erxleben, K., Treue, S., 2008. Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. J. Neurosci. 28, 8934–8944.
Yantis, S., Jonides, J., 1990. Abrupt visual onsets and selective attention: voluntary versus automatic allocation. J. Exp. Psychol. Hum. Percept. Perform. 16, 121–134.