

Evidence for an attentional priority map in inferotemporal cortex

Heiko Stemmann^{a,1} and Winrich A. Freiwald^{b,1}

^aInstitute for Brain Research & Center for Advanced Imaging, University of Bremen, D-28334 Bremen, Germany; and ^bCenter for Brain, Minds & Machines (CBMM), Laboratory of Neural Systems, The Rockefeller University, New York, NY 10065

Edited by Michael E. Goldberg, Columbia University, New York, NY, and approved October 7, 2019 (received for review December 24, 2018)

From incoming sensory information, our brains make selections according to current behavioral goals. This process, selective attention, is controlled by parietal and frontal areas. Here, we show that another brain area, posterior inferotemporal cortex (PITd), also exhibits the defining properties of attentional control. We discovered this area with functional magnetic resonance imaging (fMRI) during an attentive motion discrimination task. Single-cell recordings from PITd revealed strong attentional modulation across 3 attention tasks yet no tuning to task-relevant stimulus features, like motion direction or color. Instead, PITd neurons closely tracked the subject's attention state and predicted upcoming errors of attentional selection. Furthermore, artificial electrical PITd stimulation controlled the location of attentional selection without altering feature discrimination. These are the defining properties of a feature-blind priority map encoding the locus of attention. Together, these results suggest area PITd, located strategically to gather information about object properties, as an attentional priority map.

attention | electrophysiology | cognition

Our brains are not passive analyzers of sensory information. Rather, they select important pieces of information at the expense of currently irrelevant ones (1). This active process, selective attention, constitutes a critical link between sensory processing and internal cognitive set. It is widely accepted, based on a wealth of data from human neuropsychology and imaging as well as nonhuman primate electrophysiology, that the focus of endogenous attention is controlled by a network of areas in parietal and prefrontal cortex (2–6). In contrast, regions of the occipital and temporal lobe are thought to support the detailed processing of visual object information. When these regions are modulated by attention (7–9), this is thought to result from top-down influences from prefrontal regions, like the frontal eye fields (FEFs), or parietal regions, like the lateral intraparietal (LIP) area (10–12).

During functional magnetic resonance imaging (fMRI) in macaque monkeys performing an attention-demanding motion discrimination task (Fig. 1A), we found robust attentional modulation in a range of visual areas including but not limited to motion-selective area MT and areas LIP and FEF (13) (Fig. 1B). In this task, monkeys were required to track 1 of 2 random dot surfaces (RDSs) that were rapidly changing motion direction until directions ceased changing. This prolonged motion event (PME) had to be detected, and the direction of motion had to be reported by an eye movement in the same direction (and onto 1 of 8 peripheral saccade targets [STs]) (Fig. 1A).

Because the task emphasized motion discrimination, involvement of areas with strong directional tuning, like dorsal-stream area MT (14, 15), was expected. Because the task emphasized sustained endogenous attention, involvement of attentional control areas LIP and FEF made sense as well. Yet curiously, an additional area located in posterior and dorsal inferotemporal cortex, area posterior inferotemporal cortex (PITd) (16–18) (Fig. 1B) (a ventral-stream area), stood out as the one strongly attention-modulated area known neither for motion selectivity nor for attention control. We thus set out to determine its functional properties and targeted PITd for electrophysiological recordings.

The first multiunit signal that we recorded during the attentive motion discrimination task is shown in Fig. 1C. With attention paid into the receptive field (RF), activity was high; with attention paid outside the RF, activity fell to the level of prestimulus activity. Because of strength and trial-to-trial reliability of attentional modulation, the attentional state of the animal could be predicted, with complete fidelity (receiver operator characteristics [ROC] analysis: area under curve [AUC] = 1) (Fig. 1C, *Right*; sample recordings are in *Movies S1–S3*). An isolated single unit recorded simultaneously showed a similar degree of attentional modulation (*Methods* and Fig. 1D) (attention index [AI] 0.55, multiunit 0.69). In contrast, the activity of the cell was hardly, if at all, modulated by motion direction (Fig. 1D, *Right*) (direction indices = 0.03 and 0.07 during rapid motion events and PMEs, respectively). Thus, this particular PITd site carried little information about the attended feature but a lot about the subject's attentional state.

This pattern of strong task dependence and weak direction dependence was characteristic for the population of PITd cells ($n = 190$) as a whole. The population response showed a separation of response magnitude with attention direction growing over time (Fig. 2A). Attentional modulation was strong in the entire population of cells, with the distribution of AIs shifted almost entirely to positive values with a mean AI of 0.62 in the interval 1,500 to 3,500 ms after stimulus onset. This corresponds to a 426% increase of the attended over the nonattended response (Fig. 2B). PITd neurons were thus highly informative about the attentional state of the subject (average AUC = 0.86) (Fig. 2B, *Inset*).

In contrast, directional modulation was weak in the entire population (mean direction index 0.02, corresponding to a mean

Significance

A fundamental dogma in the cognitive neurosciences is that attention is controlled by parietal and prefrontal areas. Here, we show that an area in the temporal lobe exhibits the properties of a priority map coding the focus of attention. We show this through whole-brain functional magnetic resonance imaging, electrophysiological single-unit recordings, and causal electrical stimulation. This discovery changes our understanding of the organization of visual pathways and the functions of attention networks.

Author contributions: H.S. and W.A.F. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Data deposition: Data for this article have been deposited in figshare, <https://doi.org/10.6084/m9.figshare.c.4705649.v1>.

¹To whom correspondence may be addressed. Email: stemmann@uni-bremen.de or wfreiwald@mail.rockefeller.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1821866116/-DCSupplemental.

First published November 4, 2019.

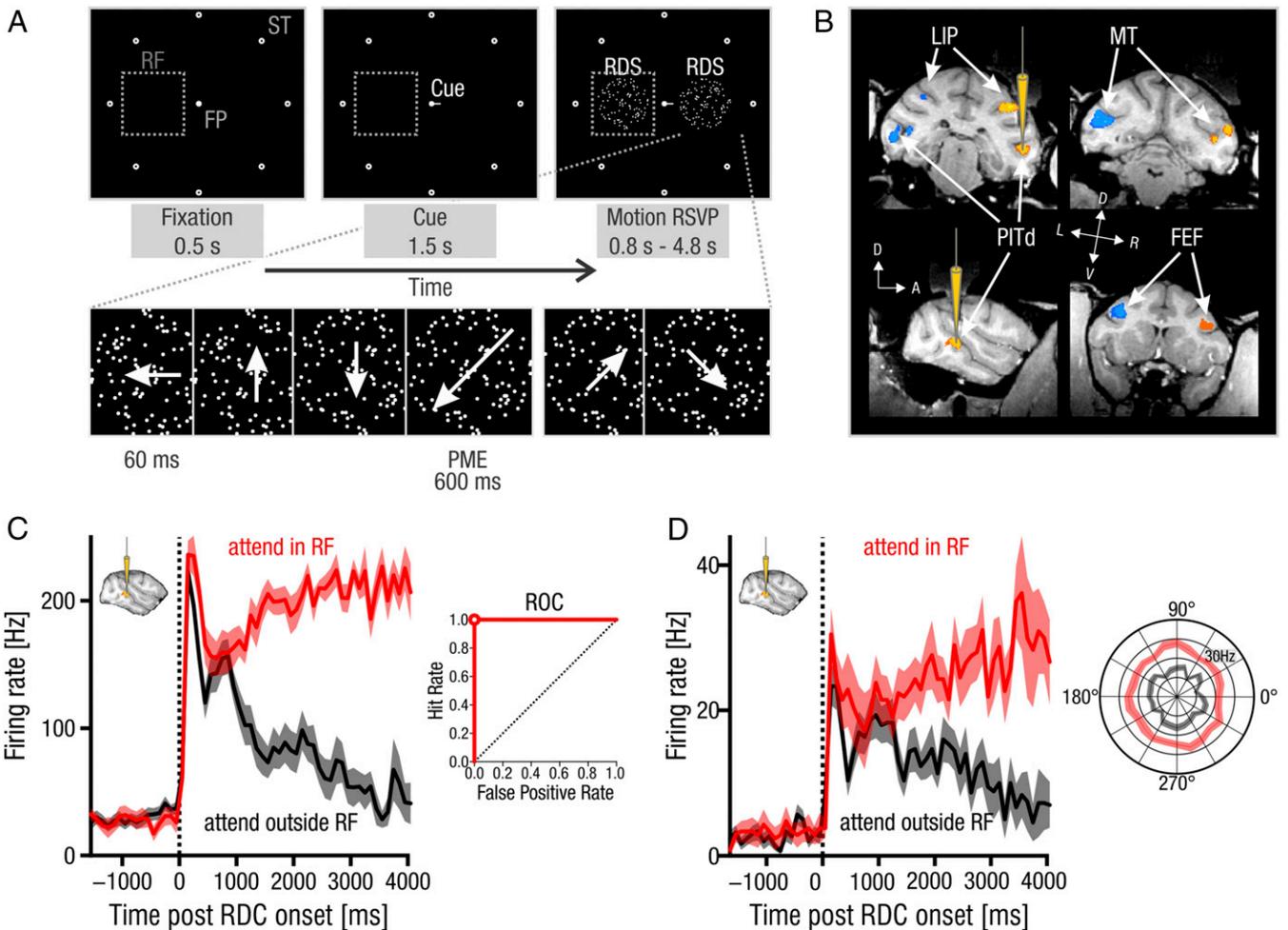


Fig. 1. Attentional modulation during a motion discrimination task. (A) Stimulus array and event sequence. Subjects initiated a trial by foveating the central fixation spot (FP) surrounded by 8 STs. After a 500-ms delay, a bar cue appeared, indicating where attention had to be deployed. After 1,500 ms, 2 RDSs appeared at opposite and equidistant positions from the fixation spot, one overlapping with the RF of the recorded cell. (Lower) While both RDSs were changing their motion direction every 60 ms, the subject had to track the target stimulus for 20 to 60 direction changes until the translation direction ceased changing for 600 ms (PME) in monkey Q and 800 ms in monkey M followed again by rapid direction changes. Monkeys were required to respond to the target surface PME by a saccade to the corresponding ST. *Methods* has details. Thus, the task emphasized sustained attention and dissociated the focus of attention from saccade planning. (B) Coronal and parasagittal slices showing the statistical parametric map of the contrast “attend contra vs. attend ipsi” (thresholds at $P < 0.05$, corrected.) in one subject’s brain through LIP, PITd, MT, and FEF. Cyan and blue indicate significantly higher activity for contrast attend left vs. right, and yellow and red indicate significantly higher activity for contrast attend right vs. attend left. The attention-modulated part of area PITd was located 4 mm anterior to the interaural line in the lower bank of the superior temporal sulcus. L, left; R, right; D, dorsal; V, ventral. (C, Left) PSTH of multiunit activity recorded in PITd (*Inset*) when attention was paid into the RF (red) or to the surface on the opposite site of the fixation point outside the RF (gray). Averages and 95% confidence levels are indicated as solid lines and transparent surroundings, respectively. Single-trial activity levels in the 2 conditions were almost nonoverlapping, reaching perfect discrimination of attentional state with an AUC level of 1 in an ROC analysis (C, Right). (D) PSTH of simultaneously recorded single unit showing similar effects with tuning curves for RSVP (rapid serial visual presentation) period during both attention conditions. PSTHs were calculated for the RSVP period of trial before the PME.

modulation by motion direction of 4%) (Fig. 2 B, Right). This lack of direction selectivity was not the result of weak visual responsiveness. In fact, PITd neurons were so highly visually responsive, it was even possible to map their RFs with a sparse white noise stimulus and quantify RF size (in 81 of 91 cells) (*Methods* and Fig. 3A). PITd RFs were spatially confined with RF sizes closely matching eccentricity (Fig. 3B). Therefore, during the attention task, a given PITd cell was driven by the RDS placed inside its RF and only minimally, if at all, by the other RDS placed equidistantly on the opposite side of the fixation spot (*Methods*). In other visual areas, strong attentional modulation occurs when both target and distractor reside inside the RF (9, 19), and this has been attributed to interstimulus competition (20). PITd, in contrast, does not require interstimulus competition for strong attention effects.

Thus, PITd neurons exhibit 1) strong attentional modulation without interstimulus competition, 2) high visual responsiveness 3) within spatially restricted RFs, and 4) little tuning to the task-relevant feature (Figs. 1, 2, and 3 A and B). These are the key properties of an attentional priority map encoding the current locus of attention (21). The attentional priority map is thought to be a processing stage that abstracts from the featural composition of stimuli and encodes instead their behavioral relevance in a spatial map. This representation can then be used to direct subsequent behavior (22–26). This role of the priority map in the control of which locations in the visual scene are attended is close to that of the “master map” postulated in the “Feature Integration Theory” (27). For an area to be considered a priority map, it must match a number of criteria: for an area to encode the location of attention, it must have spatially confined RFs tiling the visual

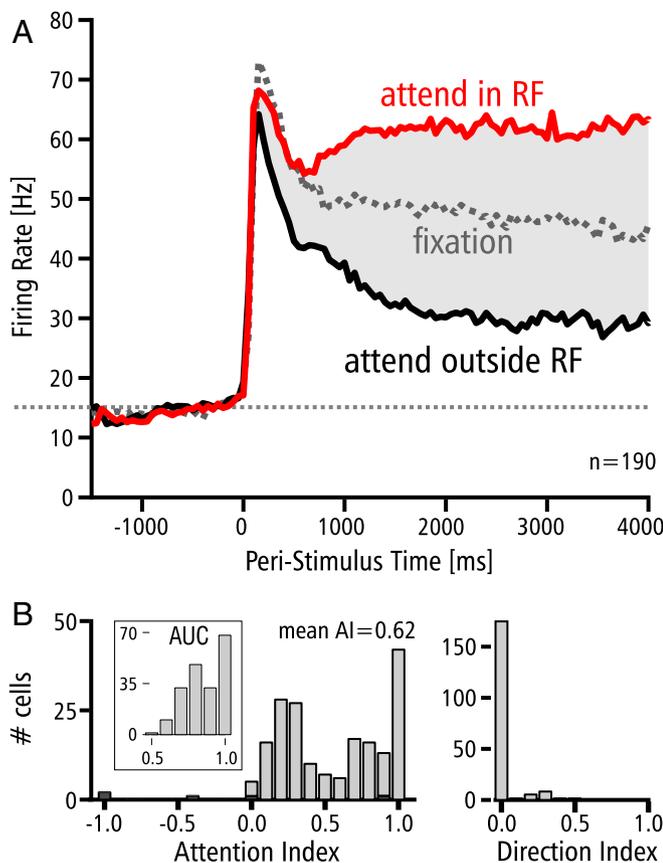


Fig. 2. PITd physiology: attention. (A) Population PSTH ($n = 190$) during motion discrimination and during passive fixation task (dotted line; otherwise conventions are the same as in Fig. 1C). With attention paid into the RF, activity stays high, while it falls off when attention is paid outside the RF. (B, Left) Histogram of AIs. Dark gray entries denote indices of cells with stimulus-induced response suppression, and light gray entries denote indices of cells with response enhancement. The distribution is shifted to the right (median AI = 0.62), with many cells showing complete attentional modulation. *Inset* shows distribution of AUC values. More than one-third of all cells are perfect indicators of the focus of attention. (B, Right) Histogram of direction-tuning indices. Almost all cells had tuning indices close to 0 (i.e., they were not tuned).

field. For the area to encode the attentional focus of an object, its cells should exhibit little feature selectivity and high visual responsiveness, thus being able to encode attention to a wide range of stimuli (21, 24). Furthermore, the area should exhibit strong attentional modulation, and ideally, the strength of this modulation should be independent of stimulus properties. Area PITd thus meets all of these functional criteria.

While highly responsive to 2 stimulus types minimal in shape (single and random dots), PITd cells might still exhibit some shape selectivity, a hallmark of inferotemporal cortex (17, 28) and the neighboring face patches (13, 29–31). This would constitute a deviation from the ideal properties of a priority map. We determined shape selectivity with 2 stimulus sets from the literature on posterior IT (inferotemporal cortex) (17, 29) (*Methods* and Fig. 3C). PITd cells were, again, very strongly modulated by stimulus presence but also, exhibited varying degrees of responsiveness to the very different shapes in the 2 stimulus sets (Fig. 3C): 57 of 96 units (60%) showed a significant modulation (Kruskal–Wallis, $P < 0.01$) for the geometric stimulus set (*Methods* and Fig. 3C, Right). On average, neurons in PITd responded to 17 (± 5) different-shaped stimuli with more than half of the maximum response elicited by the best shape stimulus in the face- and object-localizer stimulus set (96 different stimuli) and to 7 (± 2) within the geometric stimulus

set (49 different stimuli) as revealed by the pattern preference index. Stimulus preference differed widely from cell to cell, and little categorical preference was observed for the population—in stark contrast to the immediately neighboring face patches (29–31). Thus, PITd combines broad visual responsiveness with the capacity to encode a wide range of stimulus shapes.

A further requirement for a priority map is that it should be engaged similarly across stimulus dimensions and cognitive requirements (21). This is because the computational utility of the priority map relies on the generality of its involvement in attentional selection. Along this line we devised 2 additional tasks, one requiring a different kind of cognitive operation, detection, instead of discrimination and the other the processing of a different stimulus domain, color, instead of motion (*Methods* and Fig. 4A, Left and B, Left). During performance of the motion detection task, a coherent motion event (CME) in an otherwise incoherent RDS had to be detected and a saccade generated to the target stimulus. Thus, while in the discrimination tasks the focus of attention and ST were dissociated, we designed the detection task such that subjects would saccade to the focus of their attention. This feature allowed us to determine whether any of the attention effects during discrimination might get altered with motor planning. We recorded from 130 neurons in both motion discrimination and detection tasks and found strong attentional modulation in both (motion discrimination: mean AI = 0.47 ± 0.007 [mean AI monkey Q = 0.76 ± 0.01 , mean AI monkey M = 0.25 ± 0.007], motion detection: mean AI = 0.47 ± 0.008 [mean AI monkey Q = 0.83 ± 0.02 , mean monkey M = 0.21 ± 0.01]) (Fig. 4A and *SI Appendix*, Fig. S1). We further recorded from 78 neurons in both the motion and color discrimination tasks and again, found strong modulation in both (motion discrimination: mean AI = 0.76 ± 0.01 , color discrimination: mean AI = 0.75 ± 0.021) (Fig. 4B). Only a few neurons were tuned to motion direction (6.2% during the RSVP [rapid serial visual presentation] period and only 2% during the PME) or to hue (7.7% during the RSVP period and 9.0% during the PME). In 58 cells, we were able to record from all 3 tasks and found their degree of attentional modulation to be highly correlated across tasks ($r \geq 0.80$) (Fig. 4C). Thus, PITd neurons exhibit similar patterns of strong attention modulation across cognitive demands and feature dimensions.

Attention effects in PITd were so strong that it was often possible to predict online by the momentary multi- or single-unit firing rate where the subject was paying attention and thus, whether the subject was going to make a mistake. For example, when a PITd neuron's activity was high while attention was cued outside the RF, the subject would report the motion direction of the distracter inside the RF as if it had paid attention to the RF (Movies S1–S3). During these selection errors (correct reports of motion direction from the distracter), the pattern of PITd activity was reversed compared with that during correct selection (Fig. 5) as if the focus of attention was shifted to the location opposite to the cued one. However, during a different kind of error, the discrimination error (report of a motion direction present on neither target nor distracter), the pattern of PITd activity was reduced in amplitude (Fig. 5) as if now the intensity of attention was reduced. Thus, PITd activity strongly predicted both direction and quality of attentional deployment: it reflected where information was selected from—the defining property of selective attention—and it resembled the animal's attentional state.

These findings raise the possibility that PITd activity does not just track attention state closely but actually drives attentional selection. If this hypothesis is correct, then artificial activation of a PITd site should increase attentional selection from that site's RF (Fig. 6A). The final criterion, in addition to the functional characteristics discussed above, for an area to qualify as a priority map is that its artificial activation should alter attentional selection. We tested this prediction with electrical stimulation (32). If PITd is an attentional priority map, then artificial activation

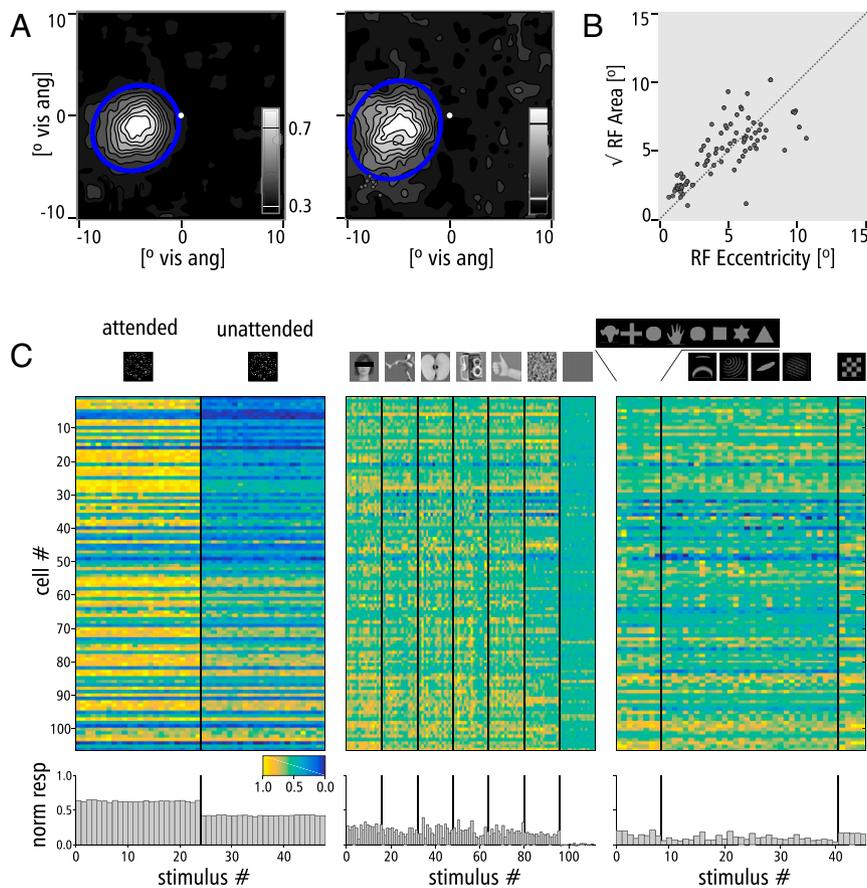


Fig. 3. PITd physiology: RFs and shape tuning. (A) Contour plots of 2 sample RF maps determined by sparse white noise mapping. Normalized activity (white: 1, black: 0) is shown as a function of dot position. Blue lines mark results of Gaussian fit (square root 2 times width and height, the area encompassing 85% of the signal). (B) Scatter plot of RF size (square root of area) and eccentricity of RF center ($n = 81$ cells). (C) Comparison of PITd cell ($n = 106$) activation across 3 tasks: (Left) attentive motion discrimination, (Center) face/object selectivity (61), and (Right) shape selectivity (17, 62). Minimum–maximum normalized activity (color coded; yellow: 1, blue: 0) is shown as a function of cell number (top to bottom; sorted by strength of attention effect) and stimulus condition (left to right)—please note that minimum–maximum normalization was applied over all 3 stimulus conditions (i.e., differences in color scale between stimulus conditions are representing differences in neural activity). Sample stimuli are shown in Top. PITd neurons are most strongly and systematically modulated by attention, not by shape. Different PITd neurons exhibit different pattern selectivity such that the mean population activity differs little across categories, including faces and scrambled patterns.

should enhance attention for stimuli within the RFs at the stimulation site and possibly reduce attention at other locations. We wanted to test the role of PITd more specifically and thus, modified the motion discrimination task slightly to enhance its sensitivity. To enhance behavioral readout sensitivity, we increased attentional load by lowering motion coherence (*Methods*). We also synchronized PME between surfaces to allow for both to be paired with electrical stimulation (occurring randomly during half the trials) (*Methods*). In each trial, of the 8 possible motion directions, one was chosen for the target, and a different one was chosen for the distractor. Thus, 4 main behavioral outcomes could occur: the subject could saccade into the motion direction displayed by the target (“hit”) or the direction of the distractor (“selection error”) or to 1 of 6 remaining STs (“discrimination error”), or the subject could fail to respond to the PMEs (“missed detection”). This set of behaviors allowed us to test, with very high granularity, the critical predictions emerging from the hypotheses that PITd is an attentional priority map.

First, if PITd is a priority map, stimulation should increase attention and thus, reduce the fraction of missed detection errors. Second, electrical stimulation in PITd should draw attention to the stimulus processed by the site of stimulation (Fig. 6A): when the target stimulus is inside the RF, its processing should be improved (increased fraction of hits); when the distractor is inside the RF, it should be erroneously selected but its motion

direction reported correctly (increased fraction of selection errors). Third, even the very strong and artificial activation of PITd should not interfere with the quality of motion discrimination (no increase in discrimination errors). These are very specific and strong predictions of the attentional priority map hypothesis.

We found the following pattern of results (Fig. 6B and C and *SI Appendix* have statistics [multinomial logistic regression] and details). First, electrical stimulation in PITd reduced the fraction of missed detection events (from 20 to 7%). Second, when the target was inside the RF, electrical stimulation increased the fraction of hits (from 70 to 81%) (Fig. 6B), and when the distractor was inside the RF, electrical stimulation increased selection errors instead (from 2 to 12%) (Fig. 6C). Third, electrical stimulation did not alter the fraction of discrimination errors (2%). Thus, electrical stimulation in PITd caused a complex profile of behavioral improvement and deterioration, and that pattern matched the predictions of the attention priority map hypothesis of PITd precisely.

This pattern of causality relating artificial PITd activation to behavior paralleled the electrophysiological profile of PITd activity (Fig. 2). First, at a time when PITd population activity did not yet differentiate very much between target and distractor (Fig. 2A), the effectiveness of electrical stimulation was highest (Fig. 6D and E): microstimulation of PITd at target location (Fig. 6D) was so effective, it decreased the fraction of missed detections from about 50 to 10%, while microstimulation of PITd at the distractor

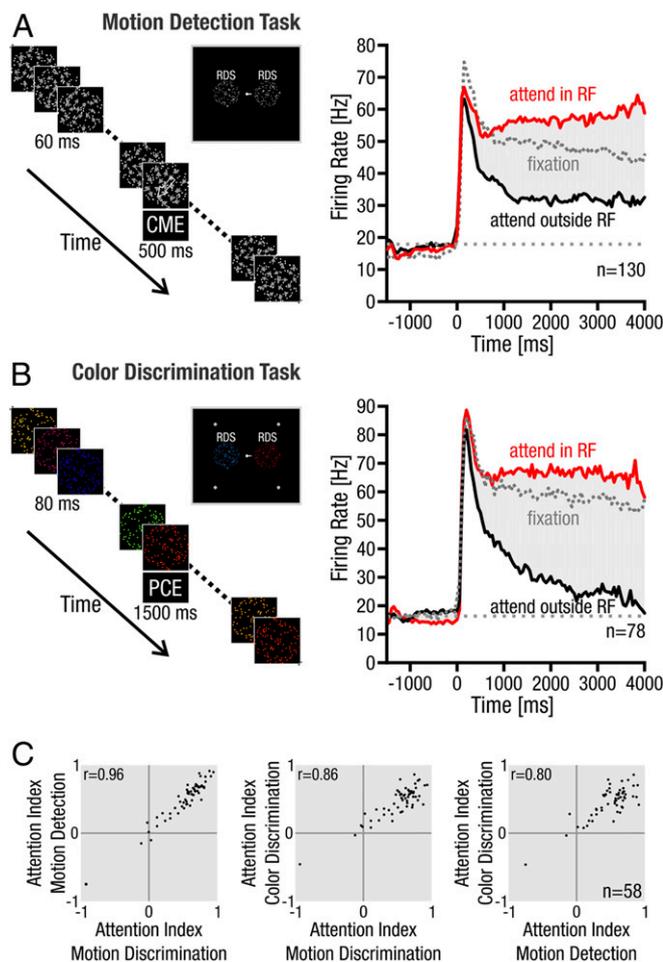


Fig. 4. Generality of attentional effects in PITd. (A, Left) Attentive motion detection task requiring subjects to detect a CME in a stream of random dot motion on target surface (Methods). (A, Right) Population PSTH of all cells ($n = 130$) tested in this paradigm. Strength and time course of attentional modulation are similar to those in the motion discrimination task (Fig. 2A). (B, Left) Attentive color discrimination task requiring discrimination of a PCE on target surface (Methods). (B, Right) Population PSTH of all cells ($n = 78$) tested in this paradigm. (C) Scatter plots of AIs of all cells ($n = 58$) measured across all 3 attention paradigms. Correlation coefficients are shown in the upper right corner. The patterns of attentional modulation were highly correlated across the 3 paradigms.

location (Fig. 6E) increased the fraction of selection errors from just above 0 to about 45% (even surpassing the fraction of hits at 38%). The effectiveness of microstimulation subsequently decreased with decay constants of about 700 to 1,000 ms (Fig. 6D and E), slightly slower than the time course of attentional differentiation in the PITd population response (Fig. 2A) ($\tau = 582$ ms). This relationship is expected when the focus of attention is determined by both natural PITd activity and the superposed artificial activation: when the former is least differentiated, the effect of the latter should be strongest, but after activity levels for target and distractor have diverged, electrical stimulation becomes ineffective. Second, effectiveness of microstimulation (inducing selection errors) and strength of attentional modulation (AI) correlated significantly from site to site ($r = 0.50$, $P < 0.01$, $n = 29$) (SI Appendix). Thus, the temporal and spatial profiles of attentional modulation in PITd predict its causal impact on attentional selection.

The pattern of microstimulation effects and the pattern of correlation with the physiology provide support for the hypothesis that PITd constitutes an attentional priority map with activity that

controls attention. Similarly, these patterns of results make alternative accounts implausible. We consider here the case of phosphenes, which any stimulation inside the visual system might generate. The generation of a phosphenes, which subsequently draws attention to its location, at a point in time related to the PME could explain the increase in performance and reduction of missed detection events. However, phosphenes would interfere with feature discrimination and would thus predict an increase in discrimination errors, contrary to what we observed. Furthermore, the strength of phosphenes is not expected to be correlated with the strength of attention effects that we observed. Most importantly, the time course of phosphenes effects would be the opposite of what we observed: phosphenes visibility would be lowest in the beginning of the trial, when firing rates are already high, and would increase over time as firing rates drop in the nonattended condition. Thus, the generation of phosphenes cannot explain the pattern of results that artificial stimulation of PITd generated.

Results from fMRI, electrophysiology, causal manipulation, and behavior show that an area in PITd does not serve the processing of featural detail but attentional selection. This is a function so far not associated with the temporal lobe, but usually associated with parietal and prefrontal cortex, like areas LIP and FEF, which also exhibit several characteristics of a priority map (4, 5). Why might there be a third area for attentional control, and why at such a remote location from the others? FEF and LIP both

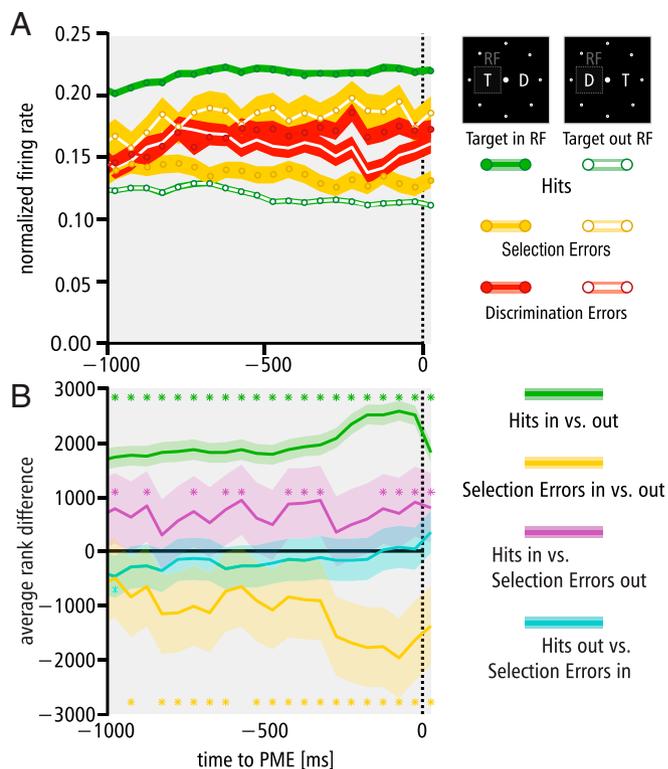


Fig. 5. PITd activity predicts behavior. (A) PITd population activity (normalized) preceding PME onset in 2 behavioral contexts (target inside or outside of RF; filled and open circles, respectively) and 3 behavioral outcomes (green: correct responses, yellow: selection errors, red: discrimination errors). PITd activity during hits (green) is higher when the target is inside the RF than outside, but the opposite is the case during selection errors (yellow). The pattern of activity during selection errors is inverted relative to the one during hits, suggesting a switch in the focus of attention, and it shows a weaker differentiation, implying reduced attentional intensity. Activity patterns during discrimination errors (red), in contrast, exhibit greatly reduced (and insignificant) differentiation. (B) Time-resolved Kruskal-Wallis test results for 4 pairwise comparisons (Right). Colored stars at the top denote significance at $P < 0.01$. T, target; D, distractor.

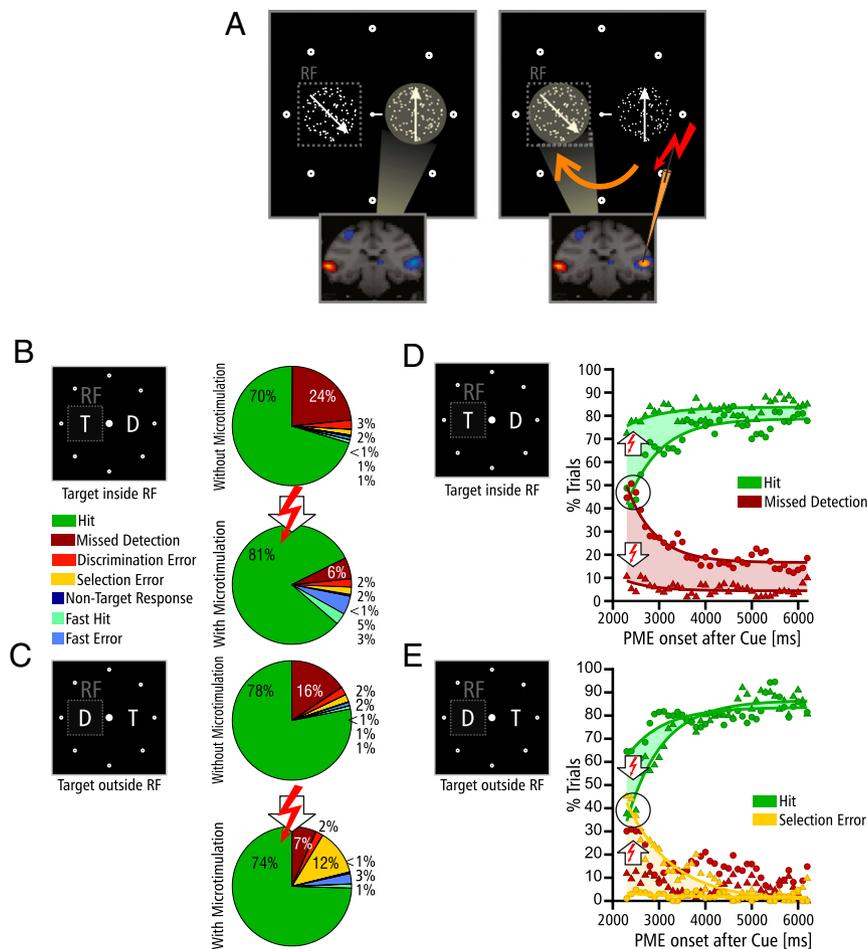


Fig. 6. Artificial activation of PITd affects behavior. (A) Schematic of microstimulation logic. If neural activity in ipsi- and contralateral PITd (Lower) determines the focus of attention (yellow), then artificial activation (Right) of an otherwise weakly active PITd could thus switch the focus of attention into that region's population RF (dotted square). (B and C) Behavioral outcomes when attention was paid inside (B) vs. outside (C) the RF without microstimulation (Upper Right) and with microstimulation (Lower Right) as pie charts. Details are in the text and *SI Appendix*. (D and E) Time courses of microstimulation effects. Color conventions are the same as in B and C (circles: data points without, triangles: data points with electrical stimulation, solid lines: exponential fits). Initially, right after stimulus onset, microstimulation effects are strongest: stimulation inside the RF improves attention to the point that otherwise prominent detection failures are almost eliminated (D), and stimulation outside the RF causes a higher fraction of selection errors than of hits (E). Effectiveness of microstimulation diminishes exponentially with time.

possess close links to oculomotor function (4, 33–35). Yet, when attention needs to be dissociated from action planning, an area devoid of these links would become important. PITd, more than LIP or FEF, is strategically positioned to gather and utilize information represented nearby on object shape (17, 28) and color (36), thus meeting a final criterion for an attentional priority map (21). Area PITd could use this property to support feature integration (27) or object-based attention (37–39).

PITd's involvement in a motion-processing task provided a puzzle (13). Motion processing is a classical function of the dorsal stream (14, 15), and thus, the involvement of PITd seemed curious (13, 40). Had we conducted a shape-processing or color discrimination task during our initial fMRI experiments, the observation of attentional modulation in PITd would not have been surprising (41, 42). An alternative account of dorsal and ventral streams, however, posits that fine feature discrimination is a function of the ventral stream (43, 44). In this framework, the involvement of PITd was less surprising but would require a representation of motion direction in PITd, which we did not find. Area PITd, we recently found, is directly connected to classical dorsal attention control areas LIP and PITd (45). The functional characteristics that we describe here for PITd show that, in many ways, it re-

sembles these dorsal-stream areas more than of neighboring temporal lobe areas, like the face patches. The findings presented here thus force a network-oriented way of thinking about neural information processing and a rethinking of old concepts about dorsal and ventral streams (14) and the interactions of parietal and temporal lobes in the control of attentional function (46–49). It is tempting to speculate in this context that area PITd might serve to relay gaze-selective signals from a nearby (or possibly even overlapping) temporal lobe area (50, 51) into parietal area LIP (52), thus mediating gaze-following behavior (52, 53) and possibly, joint attention (54). The finding of an attentional priority map in inferotemporal cortex with properties more similar to parietal and prefrontal brain regions than neighboring shape-selective areas forces a rethinking of the functional organization of the primate brain. While temporal and parietal functions are often seen as emerging in parallel through separate processing streams, PITd might be connected directly to LIP and FEF to coordinate the focus of attention, implying an orthogonal scheme of organizational that integrates functions across streams and cortical lobes. Furthermore, a lesion to the human PITd homolog (55), our results predict, would cause functional deficits other than visual agnosia, the main deficit of temporal lobe lesions, but of attentional control.

Methods

All animal procedures conformed to the National Research Council's *Guide for the Care and Use of Laboratory Animals* (56) regulations for the welfare of experimental animals issued by the Federal Government of Germany and were in accordance with the guidelines of the Caltech Institutional Animal Care and Use Committee. In brief, 2 rhesus monkeys, in which area PITd had been localized with fMRI during the performance of an attentive motion discrimination task, performed several attention and fixation tasks during electrophysiological recordings to determine basic response properties of PITd neurons and their role in selective visual attention. Recording experiments were then combined with electrical microstimulation to determine the causal role of PITd in attentive motion discrimination.

Subjects and Surgical Procedures. Two male rhesus monkeys (*Macaca mulatta*, 6 to 10 kg) were used in this study. Animals were implanted with an MR-compatible plastic head post (Ultem; General Electric Plastics) and recording chamber (Crist Instruments) attached to the skull by ceramic screws (zirconium oxide; Thomas Recording) and dental cement. All procedure followed standard anesthetic, aseptic, and postoperative treatment protocols described in detail in refs. 57 and 58.

Visual Stimulation and Tasks. Each recording day, a varnish-coated electrode (FHC Inc., 0.5- to 20-M Ω impedance) was lowered into PITd through a guide tube held in place by an fMRI-compatible recording grid (Crist Instruments). The electrode location of different guide tube positions was verified by acquiring a structural MRI before recording started. The sequence of gray and white matter passages and passages through sulci were monitored online and documented each day. All visual stimuli were generated, and behavior was controlled by a custom-made software (Visiko) running on a Windows computer system. Monkeys viewed stimuli on a CRT monitor (Iiyama HM204 DT A, 22 inches, eye-screen distance 83 cm) with a refresh rate of 100 Hz. An identical monitor was used for the experimenter to control and manually determine location and size of stimuli and RFs outside the recording room. All electrophysiological data as well as eye-position data (iscan, Inc.) and behavioral markers from the presentation system were recorded with a data acquisition system (MAP; Plexon Inc.); behavioral and visual presentation data were stored in a Visiko log file.

When the target area in PITd was reached, location and size of the RF were manually identified with the help of a manually controlled white bar stimulus (59). Bar position and orientation were controlled by a computer mouse. PITd neurons responded vigorously to the bar stimulus inside their RF. After RF location was determined, stimuli for subsequent experiments were adjusted: in attention experiments, one stimulus was presented inside the RF, while the other was positioned at equal eccentricity rotated by 180° around the central fixation spot. On a typical day, data were recorded at each electrode location first for the main attention tasks (motion discrimination task I and motion detection task in an interleaved fashion; see below), if applicable for the color attention task, and subsequently, if recording stability allowed for it, for 2 different shape-tuning tasks and an automated RF mapping procedure. On average, it was possible to record from 2 to 3 different recording positions each day for an approximate duration of 40 to 50 min per site. For some recording sites, the whole set of paradigms could not be completed due to lack of recording stability or because the subject monkey chose to terminate the experiment. On days with electrical microstimulation, data from the motion discrimination task were recorded before and after the electrical stimulation session. If possible, characterization of recording/stimulation site with shape tuning and RF mapping was attempted before and after microstimulation.

In all tasks, monkeys were required to keep fixation inside a central fixation window 1.5° and 1.75° of visual angle wide for monkeys Q and M, respectively, and 2.0° and 2.75° of visual angle high for monkeys Q and M, respectively.

Motion discrimination task I. The main task, an attentive motion-tracking task, required subjects to foveate a central fixation spot (FP; 0.25° diameter) while covertly paying attention to 1 of 2 peripheral RDSs. The target surface was cued by the direction of a short bar extending (0.35° × 0.60°) from the center of the FP. One RDS was positioned inside the classical RF of a neuron mapped during a preceding fixation task, and the other was positioned at an equidistant position found by a 180° rotation around the FP. RDSs were circular apertures optimized for the size of the RF under study. Dot density of each surface was 6 dots per square degree of visual angle, and the translation velocity was 6°/s. RDS motion always occurred with 80% coherence of all dots, while the other 20% moved in randomly assigned directions. Eye position of the animals was monitored by an infrared pupil-tracking system (ETL-200; ISCAN Inc.). RDSs randomly changed motion direction every

50 to 100 ms (brief motion events) in random multiples of 15° (drawn from a flat probability distribution). RDSs stopped changing their direction for up to 500 or 800 ms (PME) in monkey Q and M, respectively, to be followed again by rapidly changing brief motion events. The PME occurred at a random point in time after at least 10 and at most 60 brief motion events independently in target and distracter RDSs. Monkeys were required to pay attention to the target motion sequence to 1) detect the occurrence of the target PME and 2) discriminate its motion direction. Monkeys had to report the direction of the PME by a saccade to 1 of 8 peripheral STs (0.2° radius annuli with a line thickness of 0.1°) positioned 10° from the fixation spot on the cardinal and diagonal axes congruent with the motion direction of the PME. A trial was completed successfully if the animal initiated a saccade response within 500 ms after target PME onset and if the saccade reached the correct ST directly in less than 500 ms afterward. When gaze left the central fixation window, the 2 RDSs were switched off immediately. Successful completion of a trial was rewarded with a drop of water or juice. Blocks of trials of active task performance (A) were interleaved with blocks of fixation trials (F) during which a fixation spot was presented on an otherwise blank screen, and monkeys were rewarded for keeping fixation, and blocks of a passive task condition (P) with an overall stimulus configuration as in A, but no target was cued and no PMEs occurred, requiring central fixation. The sequence of blocks was repetition of the sequence AFFF. Each active and passive task condition block consisted of 6 successful trials interleaved by a 10-s block of fixation.

Motion discrimination task II. For the microstimulation experiments, the motion discrimination task was changed in the following 2 ways: PMEs of target and distracter RDSs were synchronized, and motion coherences were lowered to 50% coherence. The first change served to provide 2 equivalent sources of motion information at the same time and thus, to allow for the behavioral determination of the RDS from which information had been selected (see below). The second change was made to make PME detection more difficult and thus, allow for the evaluation of the behavioral effect of microstimulation on detection performance (see below).

Motion detection task. In the motion detection task, the brief (behaviorally irrelevant) motion events of the RDSs were completely incoherent (i.e., all dots moved independently into randomly assigned motion directions). The occurrence of the CME (10 to 25% coherence, chosen to match task performance level for this task with performance for the motion discrimination task) for up to 500 ms had to be detected and reported by a saccade onto the target surface. STs were absent in this paradigm, but all other spatial layout, temporal sequence of events, and all other task requirements were identical to those of the motion discrimination task. The motion detection task was presented in an interleaved fashion together with the motion discrimination task, which resulted in a sequence of blocks with active motion discrimination (Adiscr), fixation only (F), a passive discrimination task condition (Pdiscr), and an active motion detection (Adet) condition. Each Adiscr, Adet, and Pdiscr block consisted of 6 successfully completed trials, and each fixation block was 10 s long. The sequence of blocks was a repetition of [Adiscr F Adet F Pdiscr F].

Color discrimination task. Spatial layout, temporal sequence, and overall structure of the task were similar to the motion discrimination task, but motion direction was replaced by hue as the task-relevant dimension. Dots inside the RDS were presented statically with 0 motion, and dot density was 6 dots per square degree of visual angle. The hue of the dots changed every 80 ms until the color stopped changing for up to 1,500 ms (prolonged color event [PCE]) to be followed by rapid color changes. The color set consisted of 24 hues, which had been selected from the CIELUV color space at uniform color-angle intervals. Colors were equiluminant (6 cd/m²). The occurrence of the PCE on the target surface had to be detected, and its color had to be discriminated. The monkey had to indicate the color by a saccade to 1 of 4 STs with the matching color. STs were filled colored circles of 0.2° radius positioned on the cardinal axes at 10° eccentricity. Blocks of active color discrimination (A) were interleaved with blocks of fixation (F) and blocks of passive task condition (P) with an overall stimulus configuration as in A, but no target was cued, and no PCE occurred, requiring central fixation. The sequence of blocks was a repetition of [AFFF]. Monkey Q was trained on the color discrimination task, and recordings were taken from his area PITd.

In addition to the attention tasks, monkeys performed different fixation tasks that served to characterize spatial and shape selectivity of the neuron under investigation. During performance of all fixation tasks, the monkey was rewarded for keeping the gaze inside a central fixation window.

RF mapping. Position and spatial extent of the RF were quantitatively assessed with a sparse white noise reverse correlation technique (60). A white spot of 0.5° diameter was shown at pseudorandom positions for 300 ms without temporal gap. Possible locations for presentation were restricted to an area of 5° × 5° up to 10° × 10° around the position of maximal responses to a manually controlled white bar (59). This procedure took ~5 to 10 min of recording time. The borders

of the hand-mapped RF were marked on a transparency on the stimulus control monitor and served in the positioning of all subsequent experiments.

Shape tuning I. The stimulus set comprised 96 gray-scale images: 16 human faces, 16 human hands, 16 human headless bodies, 16 fruits, 16 technical gadgets, and 16 noise stimuli generated by phase scrambling from the gadget images (61). Pictures (size $5^\circ \times 5^\circ$) were shown at the center of the classical RF as evaluated by hand mapping before. Each image was shown for 200 ms with an interstimulus gap of 100 ms between 2 successive stimuli up to 10 times each. These stimuli have previously been used to describe shape selectivity in the middle face patches (29), face areas located immediately adjacent to the attention-modulated part of PITd studied here (13).

Shape tuning II. The stimulus set comprised 45 abstract and diverse stimuli, primarily those used by Hikosaka (17) in the first characterization of shape selectivity in area PITd of the macaque monkey. Stimuli were a star, a triangle, a square, a shell shape, a hand, a face, a cross, a circle, 5 different checkerboard stimuli with different spatial resolution (2, 4, 8, 10, and 12 cycles per 5°), and 32 binarized Gabor patches (inspired by ref. 62) with different spatial frequencies (2, 3, and 12 cycles per 5°), 12 different orientations, and 2 different degrees of curvature (straight and curved), all black and white. Stimuli were shown at the center of the RF with a fixed size of $5^\circ \times 5^\circ$ for 200 ms, with an interstimulus interval of 100 ms. While overall image sizes were identical, the overall number of black and white pixels differed somewhat between stimuli, thus potentially explaining some of the systematic response differences between stimuli.

Electrophysiological Recordings. Electrophysiological recordings were guided by structural and functional information on the location of attention-modulated PITds in each animal following the approach described in ref. 29. In brief, statistical parametric maps of the effect of covert spatial attention directed contralaterally during the attentive motion discrimination task obtained during fMRI in each animal were computed and registered to a high-resolution T1 volume of each animal (13). The recording cylinder was then implanted at a position and with a direction allowing electrodes to be safely advanced, avoiding vessels, into attention-modulated area PITd. Extracellular recordings were conducted using single Tungsten electrodes (FHC Inc.; impedance $\sim 20\text{ M}\Omega$ at 1 kHz, advanced with a Narishige drive MO-95; Narishige Japan). Electrical activity was amplified and filtered for action potential isolation with a band-pass filter at 300 to 8,000 Hz with a Plexon Multichannel Acquisition Processor (MAP) System. Spike waveforms were extracted using combinations of amplitude-time window crossings (Plexon). Spike waveforms were reassessed offline with spike-sorting software Offline-Sorter (Plexon).

Electrical Microstimulation. Electrical microstimulation was performed following electrophysiological characterization of a given PITd site with a stimulus isolator (A365; World Precision Instruments). Sequences of bipolar pulses (cathodal first), delivered at 200 Hz, were generated with a stimulator unit (S88; Grass Technologies). Cathodal and anodal pulses (80 μA) lasted 200 μs each and were separated by a 100- μs gap. For these experiments, single Tungsten electrodes (FHC Inc.) with an impedance of $\sim 100\text{ k}\Omega$ at 1 kHz were used. A train of bipolar pulses was applied starting at 200 to 300 ms before the CME onset in monkey Q and 300 to 500 ms before CME onset in monkey M and lasted 400 to 800 ms with a frequency of 200 Hz. Trials with electrical microstimulation were interleaved with trials without electrical microstimulation in a random fashion. The fraction of trials with electrical stimulation at a given site ranged from 30 to 40%.

Data Analysis. All analyses were done in MATLAB (Mathworks) and Statistica (Dell). Behavior was logged by the custom presentation software (Visiko) into a text file saved on the computer for offline analysis. The electrophysiological data together with the eye data were recorded by the data acquisition system (Rasputin; Plexon Inc.). The clocks of both systems were aligned by sending and recording TTL (transistor–transistor logic) pulses from the presentation system to the data acquisition system, and the information was merged in a first step and checked for consistency.

RF mapping. For each stimulus position, the mean firing rate across repeated presentation was calculated using a temporal window from 50 to 150 ms after stimulus onset. Firing rates were interpolated to a rectangular grid using a radial basis function interpolation and smoothed with a Gaussian kernel (2° full

width at half maximum). We then fit a 2-dimensional Gaussian to the maps with 7 free parameters (scale, rotation, width [σ_x], height [σ_y], and offsets in x , y , and z directions). Successful fits (81 of 91 maps) were then used to determine RF eccentricity, size ($\pi \times \sigma_x \times \sigma_y$), and center distance in multiples of SDs from the fixation point. In Fig. 3A, we marked the outline of the Gaussian encompassing 85% of the signal with blue curves at square root 2 times width and height.

Face and shape selectivity. For each stimulus category (faces, hands, fruits, gadgets, scrambled, bodies), the mean firing rate across multiple presentations was calculated using a temporal window from 50 to 150 ms after stimulus onset. For each unit, a nonparametric 1-way ANOVA test was performed to assess significant modulation for stimuli in the stimulus set ($P < 0.01$). In addition, the mean firing rate for each individual image was calculated over repeated presentations.

To evaluate how selective neural responses for specific stimuli were, we adopted the pattern preference index (17) that indicates the number of patterns that evoked responses with intensities over half of that of the maximum response elicited by the best pattern.

Direction and hue tuning. For each neuron recorded in the motion discrimination task, motion direction-tuning curves were computed during active task performance as well as during the passive task condition for successfully completed trials only. For calculation of the tuning curve during presentation of the brief motion events, firing rate in a time window from 50 ms to 200 ms after each motion direction onset was examined. For the prolonged motion event (PME) neural activity in a time window starting 50 ms after PME onset until the end of the PME or until the end of fixation was considered. For each unit, a direction-tuning index was computed by subtracting the activity of the nonpreferred direction from the activity of the preferred direction and a division by the sum of the 2 (pref – nonpref)/(pref + nonpref). Color-tuning curves were calculated equivalently during performance of the color discrimination task.

Attention tasks. For each single-unit activity and each multiunit activity, peristimulus time histograms (PSTHs) were calculated separately for each attention condition (i.e., attend in, attend out passive task condition and F with a bin size of 50 ms). Als were calculated by the formula (activity attend IN – activity attend OUT)/(activity attend IN + activity attend OUT) for a time period starting 1,500 ms after stimulus onset until 3,500 ms after stimulus onset. Unless stated otherwise, PSTHs are calculated for the period of short motion events excluding the PME.

ROC analysis. In order to illustrate the ability of a binary classifier system to predict the spatial location of attention within individual trials based on the neural activity in macaque area PITd, we used an ROC analysis. Only successfully completed trials (hits) were used for this analysis. Fitting of a generalized linear model using the firing rate as predictor regression was performed on the neural responses of each unit using a binomial distribution. A time window of 320 ms was used, roughly corresponding to the average reaction time for both monkeys in this task. The response to be predicted was either “attend inside RF” or “attend outside RF” hemifield. The starting time of the analysis window was varied from 1,500 ms before the behavioral relevant PME until 600 ms after in steps of 10 ms. The resulting coefficient estimates were subsequently used to determine the fitted probabilities as scores for each trial individually. The ROC for the classification of the attention state inside the RF by logistic regression was performed, and the AUC was calculated using standard built-in scripts of Matlab.

Behavioral analysis of microstimulation experiment. Analyses and results are described in *SI Appendix*.

ACKNOWLEDGMENTS. We thank Michael Borisov for stimulus programming; Aurel Wannig, Pablo Polosecki, and Ilaria Sani for discussion; Caspar Schwiedrzik for statistical advice; Aleksandra Nadolski, Nicole Schweers, Katrin Thoss, and Ramazani Hakizimana for technical support and animal care; Doris Tsao for logistic support; and Oldenburg University's Library (BIS) for providing space for manuscript writing. This work was supported by German Ministry of Science Grant 01G00506 (Bremen Center for Advanced Imaging, CAI), National Science Foundation Grant BCS-1057006, and the New York Cell Foundation.

1. M. M. Chun, J. M. Wolfe, “Visual attention” in *Blackwell Handbook of Perception*, B. Goldstein, Ed. (Blackwell Publishers Ltd., Oxford, UK, 2001), pp. 272–310.
2. S. Kastner, L. G. Ungerleider, Mechanisms of visual attention in the human cortex. *Annu. Rev. Neurosci.* **23**, 315–341 (2000).
3. E. Wojculik, N. Kanwisher, The generality of parietal involvement in visual attention. *Neuron* **23**, 747–764 (1999).

4. T. Moore, K. M. Armstrong, M. Fallah, Visuomotor origins of covert spatial attention. *Neuron* **40**, 671–683 (2003).
5. C. L. Colby, M. E. Goldberg, Space and attention in parietal cortex. *Annu. Rev. Neurosci.* **22**, 319–349 (1999).
6. T. Moore, M. Zirnsak, Neural mechanisms of selective visual attention. *Annu. Rev. Psychol.* **68**, 47–72 (2017).

7. M. Ito, C. D. Gilbert, Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron* **22**, 593–604 (1999).
8. C. J. McAdams, J. H. R. Maunsell, Attention to both space and feature modulates neuronal responses in macaque area V4. *J. Neurophysiol.* **83**, 1751–1755 (2000).
9. J. Moran, R. Desimone, Selective attention gates visual processing in the extrastriate cortex. *Science* **229**, 782–784 (1985).
10. T. Moore, K. M. Armstrong, Selective gating of visual signals by microstimulation of frontal cortex. *Nature* **421**, 370–373 (2003).
11. Y. B. Saalmann, I. N. Pigarev, T. R. Vidyasagar, Neural mechanisms of visual attention: How top-down feedback highlights relevant locations. *Science* **316**, 1612–1615 (2007).
12. G. G. Gregoriou, S. J. Gotts, H. Zhou, R. Desimone, High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* **324**, 1207–1210 (2009).
13. H. Stemmann, W. A. Freiwald, Attentive motion discrimination recruits an area in inferotemporal cortex. *J. Neurosci.* **36**, 11918–11928 (2016).
14. D. J. Kravitz, K. S. Saleem, C. I. Baker, L. G. Ungerleider, M. Mishkin, The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* **17**, 26–49 (2013).
15. L. G. Ungerleider, M. Mishkin, “Two cortical visual systems” in *Analysis of Visual Behavior*, D. J. Ingle, M. A. Goodale, R. J. W. Mansfield, Eds. (MIT Press, Cambridge, MA, 1982), pp. 549–586.
16. D. J. Felleman, D. C. Van Essen, Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
17. K. Hikosaka, Responsiveness of neurons in the posterior inferotemporal cortex to visual patterns in the macaque monkey. *Behav. Brain Res.* **89**, 275–283 (1997).
18. K. Hikosaka, Representation of foveal visual fields in the ventral bank of the superior temporal sulcus in the posterior inferotemporal cortex of the macaque monkey. *Behav. Brain Res.* **96**, 101–113 (1998).
19. S. Treue, J. H. R. Maunsell, Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* **382**, 539–541 (1996).
20. R. Desimone, J. Duncan, Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
21. J. H. Fecteau, D. P. Munoz, Saliency, relevance, and firing: A priority map for target selection. *Trends Cogn. Sci.* **10**, 382–390 (2006).
22. C. Koch, S. Ullman, Shifts in selective visual attention: Towards the underlying neural circuitry. *Hum. Neurobiol.* **4**, 219–227 (1985).
23. E. Niebur, C. Koch, Control of selective visual attention: Modeling the “where” pathway. *Adv. Neur. In* **8**, 802–808 (1996).
24. D. K. Lee, L. Itti, C. Koch, J. Braun, Attention activates winner-take-all competition among visual filters. *Nat. Neurosci.* **2**, 375–381 (1999).
25. A. E. Ipata, A. L. Gee, J. W. Bisley, M. E. Goldberg, Neurons in the lateral intraparietal area create a priority map by the combination of disparate signals. *Exp. Brain Res.* **192**, 479–488 (2009).
26. J. W. Bisley, K. Mirpour, The neural instantiation of a priority map. *Curr. Opin. Psychol.* **29**, 108–112 (2019).
27. A. M. Treisman, G. Gelade, A feature-integration theory of attention. *Cognit. Psychol.* **12**, 97–136 (1980).
28. K. Tanaka, Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* **19**, 109–139 (1996).
29. D. Y. Tsao, W. A. Freiwald, R. B. H. Tootell, M. S. Livingstone, A cortical region consisting entirely of face-selective cells. *Science* **311**, 670–674 (2006).
30. D. Y. Tsao, S. Moeller, W. A. Freiwald, Comparing face patch systems in macaques and humans. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19514–19519 (2008).
31. E. B. Issa, J. J. DiCarlo, Precedence of the eye region in neural processing of faces. *J. Neurosci.* **32**, 16666–16682 (2012).
32. C. D. Salzman, K. H. Britten, W. T. Newsome, Cortical microstimulation influences perceptual judgements of motion direction. *Nature* **346**, 174–177 (1990).
33. P. H. Schiller, E. J. Tehovnik, Neural mechanisms underlying target selection with saccadic eye movements. *Prog. Brain Res.* **149**, 157–171 (2005).
34. D. A. Robinson, A. F. Fuchs, Eye movements evoked by stimulation of frontal eye fields. *J. Neurophysiol.* **32**, 637–648 (1969).
35. P. Mazzoni, R. M. Bracewell, S. Barash, R. A. Andersen, Motor intention activity in the macaque’s lateral intraparietal area. I. Dissociation of motor plan from sensory memory. *J. Neurophysiol.* **76**, 1439–1456 (1996).
36. B. R. Conway, S. Moeller, D. Y. Tsao, Specialized color modules in macaque extrastriate cortex. *Neuron* **56**, 560–573 (2007).
37. A. F. Kramer, T. A. Weber, S. E. Watson, Object-based attentional selection—Grouped arrays or spatially invariant representations?: Comment on vecera and Farah (1994). *J. Exp. Psychol. Gen.* **126**, 3–13 (1997).
38. J. Duncan, Selective attention and the organization of visual information. *J. Exp. Psychol. Gen.* **113**, 501–517 (1984).
39. A. Wannig, L. Stanisor, P. R. Roelfsema, Automatic spread of attentional response modulation along Gestalt criteria in primary visual cortex. *Nat. Neurosci.* **14**, 1243–1244 (2011).
40. A. R. Bogadhi, A. Bollimunta, D. A. Leopold, R. J. Krauzlis, Brain regions modulated during covert visual attention in the macaque. *Sci. Rep.* **8**, 15237 (2018).
41. N. Caspari, T. Janssens, D. Mantini, R. Vandenberghe, W. Vanduffel, Covert shifts of spatial attention in the macaque monkey. *J. Neurosci.* **35**, 7695–7714 (2015).
42. G. H. Patel et al., Functional evolution of new and expanded attention networks in humans. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 9454–9459 (2015).
43. M. A. Goodale, A. D. Milner, Separate visual pathways for perception and action. *Trends Neurosci.* **15**, 20–25 (1992).
44. G. Mirabella et al., Neurons in area V4 of the macaque translate attended visual features into behaviorally relevant categories. *Neuron* **54**, 303–318 (2007).
45. I. Sani, B. C. McPherson, H. Stemmann, F. Pestilli, W. A. Freiwald, Functionally defined white matter of the macaque monkey brain reveals a dorso-ventral attention network. *eLife* **8**, e40520 (2019).
46. S. R. Friedman-Hill, L. C. Robertson, A. Treisman, Parietal contributions to visual feature binding: Evidence from a patient with bilateral lesions. *Science* **269**, 853–855 (1995).
47. L. Robertson, A. Treisman, S. Friedman-Hill, M. Grabowecky, The interaction of spatial and object pathways: Evidence from Balint’s syndrome. *J. Cogn. Neurosci.* **9**, 295–317 (1997).
48. M. Corbetta, G. Patel, G. L. Shulman, The reorienting system of the human brain: From environment to theory of mind. *Neuron* **58**, 306–324 (2008).
49. H. O. Karnath, S. Ferber, M. Himmelbach, Spatial awareness is a function of the temporal not the posterior parietal lobe. *Nature* **411**, 950–953 (2001).
50. K. Marciniak, A. Atabaki, P. W. Dicke, P. Thier, Disparate substrates for head gaze following and face perception in the monkey superior temporal sulcus. *eLife* **3**, e03222 (2014).
51. A. Roy, S. V. Shepherd, M. L. Platt, Reversible inactivation of pSTS suppresses social gaze following in the macaque (*Macaca mulatta*). *Soc. Cogn. Affect. Neurosci.* **9**, 209–217 (2014).
52. S. V. Shepherd, J. T. Klein, R. O. Deaner, M. L. Platt, Mirroring of attention by neurons in macaque parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9489–9494 (2009).
53. R. O. Deaner, M. L. Platt, Reflexive social attention in monkeys and humans. *Curr. Biol.* **13**, 1609–1613 (2003).
54. N. J. Emery, E. N. Lorincz, D. I. Perrett, M. W. Oram, C. I. Baker, Gaze following and joint attention in rhesus monkeys (*Macaca mulatta*). *J. Comp. Psychol.* **111**, 286–293 (1997).
55. H. Kolster, R. Peeters, G. A. Orban, The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. *J. Neurosci.* **30**, 9801–9820 (2010).
56. National Research Council, *Guide for the Care and Use of Laboratory Animals* (National Academies Press, Washington, DC, ed. 8, 2011).
57. D. Wegener, W. A. Freiwald, A. K. Kreiter, The influence of sustained selective attention on stimulus selectivity in macaque visual area MT. *J. Neurosci.* **24**, 6106–6114 (2004).
58. C. Fisher, W. A. Freiwald, Contrasting specializations for facial motion within the macaque face-processing system. *Curr. Biol.* **25**, 261–266 (2015).
59. D. H. Hubel, T. N. Wiesel, Receptive fields of single neurones in the cat’s striate cortex. *J. Physiol.* **148**, 574–591 (1959).
60. M. S. Livingstone, C. C. Pack, R. T. Born, Two-dimensional substructure of MT receptive fields. *Neuron* **30**, 781–793 (2001).
61. D. Y. Tsao, W. A. Freiwald, T. A. Knutsen, J. B. Mandeville, R. B. H. Tootell, Faces and objects in macaque cerebral cortex. *Nat. Neurosci.* **6**, 989–995 (2003).
62. J. L. Gallant, J. Braun, D. C. Van Essen, Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* **259**, 100–103 (1993).