Evidence for an attentional priority map in inferotemporal cortex

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

\textsuperscript{a}Institute for Brain Research \& Center for Advanced Imaging, University of Bremen, D-28334 Bremen, Germany; and \textsuperscript{b}Center 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 triggered 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.

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 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...
modulation by motion direction of 4%) (Fig. 2B, 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 3A 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 scene.
field. For the area to encode the attentional focus of an object, its cells should exhibit little feature selectivity and high visual responsivity, 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 postselection where information was selected from 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 ≥ 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 distractor), 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 distractor), 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. 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. If this hypothesis is correct, then artificial activation of a PITd site should increase attentional selection from that site’s RF (Fig. 6A).
PITd neurons exhibit different pattern selectivity such that the mean population activity differs little across categories, including faces and scrambled patterns.

attentive motion discrimination, (– that minimum 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; differences in neural activity). Sample stimuli are shown in the signal. (– increased attentional load by lowering motion coherence (– its sensitivity. To enhance behavioral readout sensitivity, we in– modified the motion discrimination task slightly to enhance attentions. We wanted to test the role of PITd more specifically and stimulation site and possibly reduce attention at other loca– 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 PMEs 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 distracter (“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 relaying 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...
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) (τ = 582 ms). This relationship is expected when the focus of attention is determined by both natural PITd activity and the superposed artificial modulation 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 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 phosphene, which subsequently draws attention to its location, at a point in time related to the PMEs 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 phosphene effects would be the opposite of what we observed: phosphene 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...
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 resembles 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.

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.

6 of 9 | www.pnas.org/cgi/doi/10.1073/pnas.1821866116 Stemmann and Freiwald

Downloaded by guest on March 10, 2020
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 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 warm-coated electrode (FHC Inc., 0.5- to 20-MΩ impedance) was lowered into PITd through a guide tube held in place by an MR-compatible recording grid (Crist Instruments). The electrode location of different guide tube positions was verified by acquiring proton density-weighted MR images before and after surgery. The recording chamber 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 experiment 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 position for the main attention task (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. At the end of each recording session, 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 microstimulation, data from the motion discrimination task were recorded before and after microstimulation.

In all tasks, monkeys were required to keep fixation inside a central fixation window (I.5° × 1.5°) positioned 10° from the center of the FP. The target 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 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 ms (0.5 s) after a PCE, respectively chopped (50% probability) by the fixation spot on the cardinal and diagonal axes congruent with the motion direction of the PME. The 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 direction in less than 500 ms afterward. When gaze left the central fixation window, the 2 RDSs were switched of 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 APPF. Each active and passive task condition block consisted of 6 successful trials interleaved by a 10-s block of fixation.

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 the other was positioned at equal eccentricity rotated by 180° around the position of maximal responses to 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 position for the main attention task (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. At the end of each recording session, 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 microstimulation, data from the motion discrimination task were recorded before and after microstimulation.

In all tasks, monkeys were required to keep fixation inside a central fixation window (1.5° × 1.5°) 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, before and after microstimulation. If possible, characterization of the functional response during window 1.5° and 1.75° of visual angle wide for monkeys Q and M, before and after microstimulation.

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 within randomly assigned motion directions). The occurrence of the CME (10 to 25% coherence, chosen at match task performance level) increased for the motion discrimination task II. To maximize PME detection, dots of target and distractor 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, allow for the behavioral determination of the RDSs from which information had been selected (see below), while the second change made PME detection more difficult and thus, allow for the evaluation of the behavioral effects of microstimulation on detection performance (see below).

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 independent of RDSs were presented randomly with 0% coherence. To make PME detection more difficult and thus, allow for the evaluation of the behavioral effects of microstimulation on detection performance (see below).

Motion discrimination task II. The main task, an attentive motion-tracking task, required subjects to foveate a central fixation spot (FP; 0.25° diameter) while the other was positioned at equal eccentricity rotated by 180° around the position of maximal responses to 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 position for the main attention task (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. At the end of each recording session, 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 microstimulation, data from the motion discrimination task were recorded before and after microstimulation.

In all tasks, monkeys were required to keep fixation inside a central fixation window (1.5° × 1.5°) 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 coherently 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 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
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' × 5') 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 PIT 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' × 5' 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 PITd in each animal following the approach described in ref. 29. In brief, statistical parametric maps of the effect of covert spatial attention directed contra- laterally 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 ~20 MΩ 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 crosscorrelations (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 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).

For each stimulus position, the mean firing rate across repeated successful completed trials (hits) were used for this analysis. Fitting of a psychometric function to the 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 presentations 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], height [μ], and offsets in x, y, and z directions). Successful fits (81 of 91 maps) were then used to determine RF eccentricity, size (x × y × z), 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 to 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 curve 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). As 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 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 (Visioka) 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.

ACKNOWLEDGMENTS. We thank Michael Borisov for stimulus programming; Aurel Wanng, Pablo Polosecki, and Ilaria Sanì for discussion; Caspar Schwiedrzik for statistical advice; Aleksandra Nadolski, Nicole Schwers, 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.
