

Perceptual effects of optogenetic stimulation of inferior temporal cortex: hallucination or distortion?

Reza Azadi (✉ r.azadi9@gmail.com)

National Institute of Health <https://orcid.org/0000-0001-8649-8046>

Simon Bohn

University of Pennsylvania

Emily Lopez

National Institute of Health

Rosa Lafer-Sousa

National Institute of Health

Karen Wang

National Institute of Health

Arash Afraz

NIH <https://orcid.org/0000-0002-9979-3674>

Biological Sciences - Article

Keywords: perception, vision, optogenetics, inferior temporal cortex, hallucination, macaque

Posted Date: February 8th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1331186/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Local artificial stimulation in high level visual areas of the brain induces complex perceptual events. Anecdotal descriptions of these events fall into two broad categories: ‘hallucinations’, which add a consistent and specific pictorial element to the contents of perception^{1,2}, and ‘distortions’, which transform the ongoing visual perception^{1,2}. Distortions are not pictorially consistent as they vary based on the visual input³. Systematic description of such characteristics of stimulation-induced perceptual events is a necessary step for understanding how neural activity gives rise to perception, and is also critical for development of visual prosthetic devices⁴. Here, we implanted arrays of LEDs⁵ over inferior temporal cortices of macaque monkeys and trained them to detect and report short optogenetic impulses delivered to their cortices. In a series of experiments, we observed that the ability to detect cortical stimulation highly depends on the choice of images presented to the eyes and that detection of cortical stimulation is most difficult when the animal fixates on a blank screen. Local stimulation of object selective parts of the visual cortex is shown here to induce perceptual events that are easy to detect and contain a strong distortive component. The causal contribution of inferior temporal neurons to perception does not seem to be strongly predetermined, as it depends on the state of vision. These findings also open the door to expanding the scope of visual prosthetics beyond the primary visual cortex.

Full Text

Perturbation of neural activity in the visual system alters visual perception. Understanding the nature of the perceptual events induced by neural perturbations is essential for bridging the causal gap between neuronal activity and vision as a behavior. This knowledge is crucial for identifying the neural underpinnings of visual hallucinations in psychiatric disease and developing effective visual prosthetics for patients with severe visual impairment.

Verbal reports of human patients describe two different types of perceptual events induced by stimulation of various parts of the visual system; here we liberally categorize them as ‘*hallucination*’ and ‘*distortion*’ events (Figure 1.a). Hallucination happens when brain stimulation adds a specific visual element to the contents of perception. For example, stimulation of the primary visual cortex evokes perception of spots of light or dark known as phosphenes⁶. Phosphenes are assumed to have an additive nature in that they occur at retinotopically predictable locations in the visual field⁷ and their characteristics do not seem to depend on the images cast on the retinae. Similarly, stimulation of face- and color-sensitive subregions of the fusiform gyrus is reported to induce perception of ‘facephenes’ (hallucinatory faces) and ‘rainbows’ respectively, independent of the object being viewed^{1,2}. Historically, stimulation induced hallucinatory events have provided the main promise and shaped the conceptual framework for the development of visual prosthetics⁴. If stimulation of a given cortical location elicits a specific hallucinatory element (eg. a phosphenes, facephene, etc.), one can directly merge and mix these hallucinatory elements to restore a rich sense of vision^{8,9}. Nevertheless, stimulation induced ‘distortion’ events complicate the landscape. In a distortion event cortical stimulation distorts the concurrent

contents of visual perception, thus the perceptual outcome of the stimulation depends on both the cortical position as well as the visual input to the brain. For instance, electrical stimulation of human fusiform gyrus is reported to induce ‘changes’ while the subjects were looking at faces. Subjects described the effects as “Your face metamorphosed.”¹⁰, or “Middle of the eyes twist”¹¹. Notably, while the subject from one of the studies was looking at nonface objects, electrical stimulation with the same parameters elicited less notable perceptual changes¹⁰. Moreover, stimulations in the face-selective fusiform area can rarely be detected with closed eyes¹², suggesting dependence of the stimulation-induced perceptual event on the visual input. If stimulation of a single position in the cortex induces different perceptual events depending on the state of the rest of the visual system, then the straightforward mix and merge approach for making visual prosthetic devices becomes obsolete, and our understanding of how the visual cortical activity is decoded by the rest of the brain needs to be revised.

While anecdotal human reports provide invaluable insights, high throughput and systematic study of the case is impossible without the appeal to non-human primate research. Phosphenes have been reliably replicated and studied in the primary visual cortex of macaque monkeys¹³. Electrical stimulation of macaque middle temporal (MT) cortex, biases the animals’ perceptual judgments in direction detection¹⁴, and depth discrimination tasks¹⁵. Cortical stimulation of disparity-tuned neurons in area V4¹⁶ and the inferior temporal (IT) cortex¹⁷ biases depth perception. Stimulation of IT gloss-selective neurons induces corresponding biases in a gloss discrimination task¹⁸. Stimulation of face-selective subregions of IT cortex decreases the threshold of detecting faces¹⁹ and optogenetic silencing of small clusters of face-selective neurons takes a toll on the ability to discriminate faces²⁰. Stimulation of face-selective parts of IT cortex is shown to strongly affect match-to-sample performance for faces but not other stimuli²¹, a result that is suggestive of face-specific distortions, but can be explained by face hallucinations as well because a hallucinatory face may interact with the match-to-sample task more for faces than the other stimuli. While these studies reveal specific perceptual changes resulting from artificial perturbation of the neural activity, they remain mostly agnostic with respect to the hallucinatory versus distortive nature of those changes.

In this study, we designed a novel psychophysical task to systematically investigate the characteristics of the perceptual events evoked by optogenetic stimulation of the IT cortex in two macaque monkeys (*Macaca mulatta*). Optogenetics offer enticing advantages over electrical stimulation²² but historically optogenetic studies often struggle to obtain large behavioral effects in nonhuman primates²³. The lack of robust behavioral effects may be the result of the large size of the primate brain, use of psychophysical tasks biased to prior assumptions and the constraining nature of acute stimulation preparations. Here we used the Opto-Array⁵, a novel chronically implantable array of LEDs, to stimulate the same cortical sites across many sessions. We utilized the optogenetic stimulation in the context of a sensitive stimulation-detection task²⁴⁻²⁶ unrestricted by prior assumptions about function of the stimulated neurons. We trained the animals to behaviorally detect a short optogenetic stimulation impulse delivered to their IT cortex while fixating at images of various objects and scenes (Figure 1.b). In each trial, following fixation,

an image was displayed on the screen for 1 s. In half of the trials, randomly selected, a 200 ms illumination impulse was delivered to IT cortex halfway through the image presentation, and the animal was rewarded for correctly identifying whether the trial did or did not contain cortical stimulation. The image content was independent of whether brain stimulation would or would not occur and the subjects' exclusive behavioral task was to detect if brain stimulation occurred in a given trial. We found this approach produced robust and large behavioral effects. Here we present the results of a series of experiments deploying this tactic, beginning with an experiment designed to determine whether optogenetic stimulation of IT cortex evokes a detectable visual event and culminating with a systematic test of the hallucination versus distortion hypotheses.

Opto-Arrays were implanted over the central IT cortex where we had previously injected Adeno Associated Virus 5 engineered to express excitatory opsin C1V1 (right and left hemispheres in monkeys Ph and Sp respectively; Figure 1.c). We also implanted an array in the corresponding region of IT in the opposite hemisphere where no virus was injected (control site).

Training Phase Results

The animals learned the task in only a few sessions, yet they were not able to detect cortical illumination over the control sites over the entire course of the training and Experiment 1 (catch trials). Figure 1.d shows the performance of monkey Ph as a function of session number during the training phase (monkey Sp performed similarly). The difference between stimulation report rate in stimulation and non-stimulation trials became and stayed statistically significant after only 4 and 11 sessions respectively in monkeys Ph and Sp (Figure 1.d, red and blue lines. arrow: Ph: $X^2(1, N= 1337) = 6.7, p = 0.010$ and Sp: $30.4 X^2(1, N= 1337) = 30.3, p < 0.001$). This difference stayed significant throughout the training phase (Ph: $p < 0.010$, Sp: $p < 0.001$ for all sessions). In contrast, performance on catch trials (Figure 1.d, yellow line) did not differ from non-stimulation trials (Ph: $p > 0.142$, Sp: $p > 0.054$ for all sessions) implying that on stimulation trials the animal is in fact reporting detection of cortical stimulation, rather than some other artifact of LED illumination, such as heat or light. These results show for the first time that excitatory optogenetic stimulation in IT cortex is behaviorally detectable in monkeys, but are agnostic to the character and nature of the perceptual event.

Experiment 1: the detection profile

To determine if the perceptual event evoked by optogenetic stimulation of IT is visual in nature, monkeys performed the stimulation-detection task with an imageset that consisted of 40 novel images including a condition with no image where the subject only viewed the uniform gray background (Extended Data Figure 1.a). In the stimulation trials (50% of trials), we randomly interleaved stimulations of two cortical sites (~3 mm apart). We found that the performance in detection of cortical illumination systematically varies while the animals fixate at different images creating a unique array of performances for each cortical position that we name the '*detection profile*'. Figure 2.a shows a detection profile for one stimulation site in monkey Ph (see Extended Data Figure 2 for more). Performance levels were

significantly different across the images (permutation test of randomly selected images per trial Ph: $p < 0.001$, Sp: $p < 0.001$ for both stimulation sites). The animals' performance for the no image condition was the lowest in all detection profiles obtained (permutation test of randomly selected images versus no image for each trial Ph: $p < 0.001$, Sp: $p < 0.001$ for both stimulation sites). These results suggest that the perceptual event evoked by stimulation is visual as its detectability interacts with the visual input.

The detection profiles of the two neighboring stimulated sites in each animal were correlated (Pearson's $r(39) = 0.91$ and 0.82 respectively in Ph and SP; $p < 0.001$ for both subjects) yet significantly different from each other (Figure 2.b for Ph and Extended Data Figure 3 for Sp). The correlation may result from leakage of light and/or shared neural resources as the two sites were only ~ 3 mm apart. A Pearson's correlation analysis of the hit rates derived from two distinct stimulation sites revealed that the correlation between patterns of performance across the imageset was significantly larger within a stimulation site than across sites (Ph: $p = 0.009$, Sp: $p = 0.002$). This difference indicates that the detection profile changes with cortical position and doesn't reflect potential image specific variations in attention. Detection profiles were uncorrelated across the two subjects (Pearson's $r(36) > 0.07$ and < 0.29 , $p > 0.077$ for all four comparisons).

Experiment 2: stability of the image effects

Experiment 1 shows that looking at some images helps the animal detect IT stimulation and that the rank of images for doing so varies across cortical sites and animals. The ability to detect cortical stimulation also likely depends on nonspecific factors such as virus expression heterogeneity and potential tissue build up under the array that may vary the effective cortical illumination power. In Experiment 2 we tested whether the rank order of the images remained constant across different illumination powers (7 different powers and 5 visual stimuli including no image condition). Figure 2.c shows the psychometric functions obtained from one stimulation site in Experiment 2 for Monkey Ph (see Extended Data Figure 4 for more). While illumination had a significant effect on performance (Pearson's correlation: Ph: $r(33) > 0.78$, $p < 0.001$, Sp: $r(53) > 0.79$, $p < 0.001$ for both stimulation sites), the choice of image did so as well (permutation test of randomly selected images per trial Ph: $p < 0.001$, Sp: $p < 0.001$ for both stimulation sites). The no image condition led to the poorest performance in both animals (permutation test of randomly selected images versus no image per trial Ph: $p < 0.001$, Sp: $p < 0.001$ for both stimulation sites).

This reveals the robustness of the effect of the visual input in detecting cortical stimulation. While we have tried our best to ensure homogeneous expression of the virus across the tissue as well as the near perfect alignment of the array with the cortex, the results of Experiment 2 make it difficult to explain the variation of image rank across cortical positions by the nonspecific factors that may influence effective illumination.

Experiment 3: hallucination vs distortion

The results of Experiments 1 and 2 can be explained by variance in the magnitude or quality of perceptual distortions across images. This interpretation implies that stimulation of a given position in IT cortex leads to various distortions for different visual inputs. Alternatively, these results can be explained with the hallucination hypothesis. According to this interpretation, cortical stimulation adds a constant hallucinatory element to the contents of perception but perceptual detection of this element varies in difficulty due to various figure-ground interactions (e.g. crowding effect) with screen images²⁷. However, performance for the no image condition remained systematically low in both experiments and all tests. This suggests that subjects actually use the screen images to detect the cortical stimulation and has important implications which can hardly be explained with the hallucination hypothesis. But the no image condition was an unusual condition among multiple visual stimuli and odd-ball psychophysical effects may have contaminated this finding, a shortcoming that inspired Experiment 3.

In order to tease apart these two interpretations, we tested how attenuation of visibility of screen images affect detection of the cortical event. The animals performed the stimulation-detection task while fixating on randomly presented images of five objects at four visibility levels in addition to a no image condition (Extended Data Figure 1). Visibility was degraded by reducing the saturation, spatial frequency and contrast of the images. Hypothetically, if the stimulation induces only a hallucinatory percept, decreasing the visibility of the screen image should either not affect detectability of the brain event or aid detection by decreasing background clutter. On the other hand, if cortical stimulation causes a distortive effect, it would be easier to notice when the screen image is more visible because the perceptual effect would be a function of the visual input (Figure 3.a).

Visibility of the image had a strong effect on stimulation detectability (Figure 3.b and Extended Data Figure 5; one-way ANOVA Ph: $F(4,16) = 5.24, p = 0.006$; Sp: $F(4,16) = 22.33, p < 0.001$). Spearman's Ph: $r(20) = 0.71, p < 0.001$; Sp: $r(20) = 0.78, p < 0.001$). Consistent with the distortion hypothesis, the monkeys' performance increased with the visibility of the visual input; the fully visible stimuli had significantly higher performance compared with the two lower levels of visibilities and the no image condition ($p < 0.001$ and < 0.001 for all comparisons for monkeys Ph and Sp). These results are consistent with the idea that IT stimulation has a strong distortive effect in perception of visual objects. Although this should be taken carefully as the animals still performed significantly above chance for the uniform gray condition. This might be because of a hallucinatory component that we cannot rule out or the distortion of the screen itself or both. Further investigations will hopefully push this frontier in the future.

These results reveal that optogenetic stimulation of a ~1 mm³ subregion of IT cortex evokes visual events that are easily detectable by the subject. These events are strongly and selectively enhanced by concurrent object related activity in the visual system and are not psychophysically isolated from the ongoing visual perception as in strong hallucination. Stimulation of a given cortical position with similar physical parameters appears to induce different perceptual distortions depending on the visual input as its detection varies with the choice of the image and depends on the visibility of that image.

The current Opto-Array technology doesn't allow recording of the neural activity, limiting the scope of this study to phenomenology of the stimulation-induced events. Yet it is hard not to speculate about the neural underpinnings of the observed effects. Stimulation of ~1 cubic millimeter of tissue⁵ is expected to engage IT cortex at a scale that still preserves object category selectivity^{28,29}. The fact that local cortical perturbation of this scale interacts differentially with various objects is consistent with the heterogeneity of object responses across IT cortex. It may be possible to explain the observed perceptual effects by feedforward models that incorporate object selective excitatory/inhibitory inputs to the targeted neurons, thus varying their population sensitivity to the artificial stimulation in the presence of different visual objects. In fact, 'hallucinations' models based on deep convolutional neural networks are very consistent with our findings in that they are structured based on the visual input^{30,31}. Note that the term hallucination is used in its inclusive definition in these studies. Alternatively, empirical neural recording data may require more complex models that incorporate feedback and cortico-cortical dynamics in order to explain how a complex system responds to local perturbation³². Further modeling and physiology work is required to understand the underlying neural mechanics of a psychophysical landscape that looks promising.

These findings invite an investigation of state dependence of phosphenes in primary visual cortex as they might not be as hallucinatory as we have assumed. As for IT neurons, their contribution to perception does not seem to be predetermined by their selectivity profile or category labels and the machinery that reads out IT neural activity seems to interpret it with respect to the global state of the visual system. The distortive nature of the perceptual events observed here alludes to the possibility of using stimulation of IT cortex in visual prosthetics to shape and perceptually organize arrays of phosphenes induced by stimulation of primary visual cortex. All together, we hope this study sparks interest and provides a new approach in the search for a causal framework that explains how the neural state shapes and constrains the phenomenal perceptual state of vision.

References

1. Schalk, G. *et al.* Facephenes and rainbows: Causal evidence for functional and anatomical specificity of face and color processing in the human brain. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 12285–12290 (2017).
2. Jonas, J. *et al.* Self-face hallucination evoked by electrical stimulation of the human brain. *Neurology* **83**, 336–338 (2014).
3. Parvizi, J. *et al.* Electrical stimulation of human fusiform face-selective regions distorts face perception. *J. Neurosci.* **32**, 14915–14920 (2012).
4. Fernández, E. *et al.* Visual percepts evoked with an intracortical 96-channel microelectrode array inserted in human occipital cortex. *J. Clin. Invest.* **131**, (2021).

5. Rajalingham, R. *et al.* Chronically implantable LED arrays for behavioral optogenetics in primates. *Nat. Methods* **18**, 1112–1116 (2021).
6. Foerster, O. Beitrage zur pathophysiologie der sehbahn und der sehsphare. *J. Psychol. Neurol., Lpz.* **39**, 463 (1929).
7. Brindley, G. S. & Lewin, W. S. The sensations produced by electrical stimulation of the visual cortex. *J. Physiol.* **196**, 479–493 (1968).
8. Dobelle, W. H., Mladejovsky, M. G. & Girvin, J. P. Artifical vision for the blind: electrical stimulation of visual cortex offers hope for a functional prosthesis. *Science* **183**, 440–444 (1974).
9. Bosking, W. H., Beauchamp, M. S. & Yoshor, D. Electrical Stimulation of Visual Cortex: Relevance for the Development of Visual Cortical Prosthetics. *Annu Rev Vis Sci* **3**, 141–166 (2017).
10. Parvizi, J. *et al.* Electrical stimulation of human fusiform face-selective regions distorts face perception. *J. Neurosci.* **32**, 14915–14920 (2012).
11. Rangarajan, V. *et al.* Electrical stimulation of the left and right human fusiform gyrus causes different effects in conscious face perception. *J. Neurosci.* **34**, 12828–12836 (2014).
12. Murphrey, D. K., Maunsell, J. H. R., Beauchamp, M. S. & Yoshor, D. Perceiving electrical stimulation of identified human visual areas. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 5389–5393 (2009).
13. Schiller, P. H., Slocum, W. M., Kwak, M. C., Kendall, G. L. & Tehovnik, E. J. New methods devised specify the size and color of the spots monkeys see when striate cortex (area V1) is electrically stimulated. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 17809–17814 (2011).
14. Salzman, C. D., Britten, K. H. & Newsome, W. T. Cortical microstimulation influences perceptual judgements of motion direction. *Nature* **346**, 174–177 (1990).
15. DeAngelis, G. C., Cumming, B. G. & Newsome, W. T. Cortical area MT and the perception of stereoscopic depth. *Nature* vol. 394 677–680 (1998).
16. Shiozaki, H. M., Tanabe, S., Doi, T. & Fujita, I. Neural activity in cortical area V4 underlies fine disparity discrimination. *J. Neurosci.* **32**, 3830–3841 (2012).
17. Verhoef, B.-E., Vogels, R. & Janssen, P. Inferotemporal cortex subserves three-dimensional structure categorization. *Neuron* **73**, 171–182 (2012).
18. Baba, M., Nishio, A. & Komatsu, H. Relationship Between the Activities of Gloss-Selective Neurons in the Macaque Inferior Temporal Cortex and the Gloss Discrimination Behavior of the Monkey. *Cereb Cortex Commun* **2**, tgab011 (2021).

19. Afraz, S.-R., Kiani, R. & Esteky, H. Microstimulation of inferotemporal cortex influences face categorization. *Nature* **442**, 692–695 (2006).
20. Afraz, A., Boyden, E. S. & DiCarlo, J. J. Optogenetic and pharmacological suppression of spatial clusters of face neurons reveal their causal role in face gender discrimination. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6730–6735 (2015).
21. Moeller, S., Crapse, T., Chang, L. & Tsao, D. Y. The effect of face patch microstimulation on perception of faces and objects. *Nat. Neurosci.* **20**, 743–752 (2017).
22. Roy, A. *et al.* Optogenetic spatial and temporal control of cortical circuits on a columnar scale. *J. Neurophysiol.* **115**, 1043–1062 (2016).
23. Tremblay, S. *et al.* An Open Resource for Non-human Primate Optogenetics. *Neuron* **108**, 1075–1090.e6 (2020).
24. Murphrey, D. K. & Maunsell, J. H. R. Behavioral detection of electrical microstimulation in different cortical visual areas. *Curr. Biol.* **17**, 862–867 (2007).
25. May, T. *et al.* Detection of optogenetic stimulation in somatosensory cortex by non-human primates—towards artificial tactile sensation. *PLoS One* **9**, e114529 (2014).
26. Dai, J., Brooks, D. I. & Sheinberg, D. L. Optogenetic and electrical microstimulation systematically bias visuospatial choice in primates. *Curr. Biol.* **24**, 63–69 (2014).
27. Toet, A. & Levi, D. M. The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Res.* **32**, 1349–1357 (1992).
28. Lafer-Sousa, R. & Conway, B. R. Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nat. Neurosci.* **16**, 1870–1878 (2013).
29. Sato, T. *et al.* Object representation in inferior temporal cortex is organized hierarchically in a mosaic-like structure. *J. Neurosci.* **33**, 16642–16656 (2013).
30. Suzuki, K., Roseboom, W., Schwartzman, D. J. & Seth, A. K. A Deep-Dream Virtual Reality Platform for Studying Altered Perceptual Phenomenology. *Sci. Rep.* **7**, 15982 (2017).
31. Bau, D. *et al.* Paint by Word. *arXiv [cs.CV]* (2021).
32. Jazayeri, M. & Afraz, A. Navigating the Neural Space in Search of the Neural Code. *Neuron* **93**, 1003–1014 (2017).

Methods

Surgical procedure

In this study, we performed 3 experiments and collected data from two adult male rhesus monkeys (*Macaca mulatta*), referred to as Sp and Ph. All procedures were conducted in accordance with the guidelines of the National Institute of Mental Health Animal Use and Care Committee.

In a sterile surgery under general anesthesia we performed a craniotomy and a durotomy to access the surface of IT cortex (left hemisphere in Sp, right hemisphere in Ph). We then injected AAV5-C1V1(t/t)-EYFP (nominal titer: 8×10^{12} particles/ml) into the cortex. To ensure uniform viral expression and reduce anesthesia-controlled time, we used an injection array¹ including four 31-gauge needles arranged in a 2×2 mm square. We placed the injection array four times, tiling central IT cortex with sixteen evenly spaced injection sites, resulting in a region of ~ 6mm x 6mm viral expression. Each needle was connected with flexible tubing to a 100 µl Hamilton syringe, and injection was controlled by a microinjection pump (Harvard Apparatus Pump 11 Elite). At each injection site, 10 µl of virus was injected at a 0.5 µl/min rate. Ten minutes were allowed to elapse after each injection before removing the array to allow the virus to diffuse into the cortical tissue.

Several weeks later (12 and 4 weeks in Sp and Ph respectively), in a second surgery, we confirmed the virus expression and implanted an Opto-Array (Blackrock Microsystems) on the injection site. To confirm the viral expression, we used the fluorescent signature of the enhanced yellow fluorescent protein (EYFP) coexpressed with the opsin in transfected cells, by shining a 490-515 nm wavelength light (with a NIGHTSEA Dual Fluorescent Protein Flashlight) and viewing the cortex through 550 nm longpass filter-goggles (NIGHTSEA). This fluorescent signature was confirmed in Monkey Sp, but not in Monkey Ph. Therefore, in Monkey Ph, before proceeding with the array implantation, we performed a second virus injection similar to the first injection procedure (3 injection array placements, yielding 12 injection sites in a region of 4 x 6 mm; 10 µl of virus injected into each site at 0.5 µl/min rate). Then we implanted an Opto-Array over the injection sites. The Opto-Array was placed directly on the pia mater and sutured to the neighboring dura. Following this, in the same surgery, we implanted a second Opto-Array on a similar area of the IT cortex in the opposite hemisphere (control site: right hemisphere in Sp, and left hemisphere in Ph) where no virus injection was performed.

Apparatus

The experiment was carried out with the monkey head fixed, positioned 57 cm from a 27 in, 3840x2160 pixel, 60 Hz, Dell P2715Qt monitor. Fluorescent room lights were turned on to avoid dark adaptation of the retinae. This was done to minimize the possibility that the monkey would detect the light from the Opto-Array through the skull. To guard against heating cortical tissue by LED activation, temperature on the LED die was monitored by a thermistor inside the Opto-Array at the beginning of each trial and trial delivery was paused if the temperature on the LED die rose more than 3° C above the baseline temperature, and restarted once they were less than 1° C above the baseline. 3° C at the LED die translates to approximately 0.5° C temperature change on the cortical surface; this temperature management regime is detailed in Rajalingham et al. 2021². The experiment was controlled with a custom MWorks script (The MWorks Project), running on a Mac Pro 2018. Opto-Arrays were controlled by

a Blackrock LED Driver (Blackrock Microsystems) running a custom firmware version for compatibility with MWorks. Gaze was tracked with an Eyelink 1000 Plus (SR Research). Animals were water-restricted in their cages and received liquid rewards for successfully completing trials.

Behavioral task

Monkeys were trained to perform a detection task in which they were rewarded if they correctly identified whether a trial did or did not contain an optogenetic stimulation impulse. The subject started a trial by fixating on a central fixation point (black-on-white bullseye, 0.4° outer diameter and 0.2° inner diameter) for 500 ms on a gray background. Then, an image (scaled so the largest dimension spanned 8° for most images and 30° for four scenes during training and two scenes in experiment 1) appeared on the screen for 1000 ms while the animal held fixation on a central target. In half of the trials (randomly selected) 500 ms from the image onset, an LED on one of the Opto-Arrays was activated for 200ms. Then the image and central fixation point disappeared and two response targets appeared on the vertical midline (white, 0.4° diameter, 5° above and below center). The subject reported the existence of cortical stimulation by fixating for 100 ms on one of the response targets. Then, the response targets disappeared and a unique sound was played for correct and incorrect responses. The subject received a juice reward for a correct response or a punishment of 3.5 s delay before starting the next trial in case of an incorrect response. Trials with broken fixations or a latency of more than 3 s for choosing a response target were considered as an incorrect response during the experiment but excluded from further analysis. A ~300 ms tone played at the same time the image appeared to indicate that a trial had started.

Throughout the training phase and all the experiments, 50% of the trials were '*no-stimulation*'. The other 50% were trials in which an opto-array was activated. In '*stimulation*' trials (40%-50% of all trials depending on the experiment and monkey, see experimental conditions for details), the opto-array on the virus-expressed site was activated and in '*catch*' trials (0%-10% of all trials) the opto-array on the control site was activated. The catch trials used the same stimulation parameters and were rewarded the same as stimulation trials. Performance above chance level on the catch trials would indicate that the subjects did not truly perform the task by detecting the optogenetic activation of IT neurons. This controlled for the possibility that the subjects might be glimpsing light through the skull, or be detecting a potential perturbation of the neural activities caused by the heat³.

Behavioral Training

Both monkeys were operantly trained on the experimental task using a different set of images than would be used in the subsequent experiments (Extended Data Figure 1b). To maximize the signal that the monkey was learning to detect, we began training by activating five LEDs simultaneously with power of 10.6 mW and 12.1 mW per LED for Ph and Sp respectively in stimulation trials. To reduce choice bias, we employed a 'correction loop' procedure⁴. Under this protocol, if the monkey chose the same wrong response target more than three times in a row, every subsequently presented trial would be the opposite type until the monkey selected the correct response target. Data collected in correction loops were

excluded from analysis. Ph. started the training phase with 2 images and the number of images was gradually increased to 22. Sp. started training with all 22 images, but we eventually reduced the number of images to 1 and slowly reintroduced the full training set like in Ph. Then, in both monkeys we reduced the number of activated LEDs to one, and illumination power to 4.5 mW in Ph and 9.1 mW in Sp. We introduced catch trials to Ph. after 17 sessions at an initial rate of 5% of all trials, then after 23 sessions increased the rate to 10% of all trials which continued for the rest of training. Catch trials comprising 10% of all trials were included for Sp. in all training sessions. In total, the subjects performed 42 and 48 sessions in the training phase, with 67,115 trials and 41,409 trials respectively for Ph and Sp. Part of the training data is reported in Rajalingham et al., 2021².

Experimental conditions and visual stimulus

Experiment 1 contained 40 images and 2 illumination sites for stimulation trials (see Extended Data Figure 1a for image set and inset in Figure 2b and Extended Data Figure 3 for schematic of illumination locations) with illumination power of 3.6 mW and 5.4 mW respectively for Ph and Sp. Catch trials were included at a rate of 10% and 2%, and 10 and 13 sessions were performed with a total of 17,033 trials and 16,125 trials, and an overall performance of 84.6% and 84.9% correct (catch trials excluded), respectively for Ph and Sp. Ph only received catch trials to one site on the control array while Sp received catch trials to two sites, randomly interleaved. The performances for detecting cortical stimulation were statistically significant for stimulation trials (Ph: $\chi^2(1, N= 15320) = 7295.1, p < 0.001$ and Sp: $\chi^2(1, N= 15794) = 7714.7, p < 0.001$) but not for catch trials (Ph: $\chi^2(1, N= 10370) = 0.02, p = 0.879$ and Sp: $\chi^2(1, N= 8428) = 1.7, p = 0.190$).

Experiment 2 contained 5 images, 2 stimulation sites and 7 intensity conditions. The stimulation sites were the same as in experiment 1. The images used in this experiment were a subset of the images used in experiment 1, with the two highest and two lowest d' image conditions selected (average of the two cortical locations). The fifth image was chosen by calculating which image had the greatest difference in d' between cortical location conditions in experiment 1. Illumination power for “stimulation” trials ranged from 0.4 mW to 5.4 mW for both monkeys and 9 and 12 sessions were performed with 14,941 and 14,056 trials collected with overall performance of 79.6% and 74.7% correct, respectively for Ph and Sp. This experiment included no catch trials.

Experiment 3 contained 5 images and 4 image visibility conditions, plus one “no image” (uniform gray) condition (see Extended Data Figure 1c for this image set). The “no image” condition occurred as often as any one “image at a visibility” condition, creating 21 total conditions. One cortical site was used for this experiment (Site 1 for both monkeys). We selected the top 5 highest d' images from experiment 1 at that cortical site for this imageset and degraded their visibility by reducing their contrast, saturation, and spatial frequency to near gray. To do this, the mean luminance of each image was adjusted to match that of the gray display background. Then, saturation was reduced by multiplying each pixel's chromaticity coordinates (a* and b*, CIELAB 1976) by a scale factor of 1/9, and 1/27 for the decreasing visibility levels. Image contrast was reduced by the same operation on the L* dimension (lightness) of the CIELAB

color profile, but first the mean L^* of the distribution was subtracted from each pixel, then re-added after multiplication by the scale factor, ensuring that the mean luminance of the distribution was unchanged. Finally, the spatial frequency of the Lab-scaled images was reduced by convolving each image with a 2D gaussian smoothing kernel with standard deviations of 0.39, 0.78, and 1.56° for the different visibility level. To ensure that the filtered images blended evenly into the background, padding was added to the edges of the images but care was taken to ensure the presented size was the same 8° as experiment 1 and 2. Each visibility condition was a combination of one CIELAB scaling factor and one gaussian filter. Illumination power was 3.5 mW and 5.4 mW, and 1 session was performed for each monkey with 2030 and 3193 trials collected with overall performance of 77.8% and 90.9% correct trials, respectively for Ph and Sp.

Data Analysis

Detection performance: we used d' as a bias-free measure of performance for detecting cortical stimulation⁵ which is estimated by the following equation:

$$d' = Z(H) - Z(F)$$

Where Z is Z-transform, H is the animal's hit rate for detecting stimulation, and F is the false alarm rate representing the proportion of trials where no stimulation was applied but the animals reported the trial as stimulated.

Effect of image on detectability (Experiment 1): first, we calculated a d' for each image, indicating the detectability of cortical stimulation. This creates a 'detection profile' shown in Figure 2.a and Extended Data Figure 1. The 95% confidence intervals are estimated for each image by bootstrapping the data, resampling 10,000 times with replacement⁶ and the violin plots represent the distribution of the bootstrapped data. To statistically test the effect of image on detectability of cortical stimulation, we ran a permutation test in which first we calculated the standard deviation of the d' s across images (observed standard deviation). Then, we generated the null distribution by randomly assigning the images to the trials with 10,000 replications and compared the observed standard deviation to the distribution of standard deviations generated from the null model. The permutation tests showed that the effect of images on detection of cortical stimulation is statistically significant ($p < 0.001$ for all the detection profiles). Moreover, we ran the same permutation tests after excluding the no image trials from the data and the result remained statistically significant ($p < 0.001$ for all the detection profiles).

Effect of cortical stimulation location on image detection profile (Experiment 1): we used Pearson's correlation to evaluate the similarity between the detection profiles derived from two neighboring stimulation sites. First we calculated a hit rate for each image and each stimulation site. The correlation between hit rate profiles at neighboring sites were statistically significant (Pearson's $r(39) = 0.91$ and 0.82 respectively in Ph and SP; $p < 0.001$ for both subjects). To determine if there was a difference between the sites, we followed up these results with bootstrapped estimates of the correlations within each site and

between them, resampled 10,000 times with replacement. The median correlation coefficients were 0.95 and 0.95 within the sites, and 0.86 between the sites for Ph (Figure 2.b) and 0.89 and 0.95 within and 0.75 between for Sp (Extended Data Figure 3). These results show that the detection profiles are more correlated within the sites compared with between them in both subjects. To test if this is a statistically significant difference, we generated a null distribution by randomly assigning sites to the stimulation trials with 10,000 replications; the results of this permutation test show that the observed correlation between the sites is smaller than the correlation between sites in the null distribution derived by randomly assigning sites to the trials (Ph: $p = < 0.010$, Sp: $p < 0.001$; Figure 2.b and Extended Data Figure 3).

Effect of Illumination power on image detection profile (Experiment 2): in Figure 2.c and Extended Data Figure 4, we plotted d 's as a function of illumination power for each image. We used the following formula to fit the data:

$$d'_i = \alpha x_i^{\beta + \beta_1 \lambda_i^1 + \beta_2 \lambda_i^2 + \dots + \beta_5 \lambda_i^5}$$

Where i is the trial number. x is illumination power, $\alpha, \beta, \beta_1, \dots, \beta_n$ are the fit coefficients. For each trial, the λ that matches the image index is assigned 1, and the rest are assigned 0. Therefore, $\beta_1, \beta_2, \dots, \beta_5$ represent the effect of the image on the psychometric functions. The range of r^2 s are from 0.89 to 0.98 and 0.83 to 0.95 with the average of 0.94 and 0.89 for both stimulation sites respectively for Ph and Sp. Then we calculated the standard deviation of the coefficients $\beta_1, \beta_2, \dots, \beta_5$. A permutation test was performed by assigning random image indices to the trials (10,000 times repetitions) to generate the null distribution of standard deviation for these coefficients. The results showed a significant effect of images on psychometric functions ($p < 0.001$ for both sites on both subjects).

Data availability

The data that support the findings of this study are available on request from the corresponding author.

Code availability

Code for experimental design and data analysis are available on request from the corresponding author.

References

1. Fredericks, J. M. *et al.* Methods for mechanical delivery of viral vectors into rhesus monkey brain. *J. Neurosci. Methods* **339**, 108730 (2020).
2. Rajalingham, R. *et al.* Chronically implantable LED arrays for behavioral optogenetics in primates. *Nat. Methods* **18**, 1112–1116 (2021).

3. Owen, S. F., Liu, M. H. & Kreitzer, A. C. Thermal constraints on in vivo optogenetic manipulations. *Nat. Neurosci.* **22**, 1061–1065 (2019).
4. Salzman, C. D., Murasugi, C. M., Britten, K. H. & Newsome, W. T. Microstimulation in visual area MT: effects on direction discrimination performance. *J. Neurosci.* **12**, 2331–2355 (1992).
5. Green, D. M. & Swets, J. A. *Signal detection theory and psychophysics*. vol. 1 (Wiley New York, 1966).
6. Efron, B. Bootstrap Methods: Another Look at the Jackknife. in *Breakthroughs in Statistics: Methodology and Distribution* (eds. Kotz, S. & Johnson, N. L.) 569–593 (Springer New York, 1992).

Declarations

Acknowledgements:

We thank Mark Eldridge, Nanami Miyazaki, and Emilia Jaskot for their critical help in the injection surgeries; Elia Shahbazi and Timothy Ma for their help with designing the implants; and Chris Stawarz and Andrew Mitz for their help with Opto-Array setup. This research was supported by the Intramural Research Program of the NIMH ZIAMH002958 (to A.A.).

Author contributions:

S.B., R.A., E.L. and R.L. designed research with guidance from A.A; S.B., E.L., R.A., R.L. and K.W. performed research; R.A., S.B. and E.L. analyzed data; R.A., S.B. and R.L. prepared figures; R.A., S.B., R.L., and A.A. wrote the manuscript; all authors reviewed the manuscript.

Ethics:

All procedures conducted in this study were approved and oversight by the National Institute of Mental Health Animal Use and Care Committee.

Competing interests:

The authors declare that there is no conflict of interest.

Additional information:

Correspondence and requests for materials should be addressed to R. A.

Figures

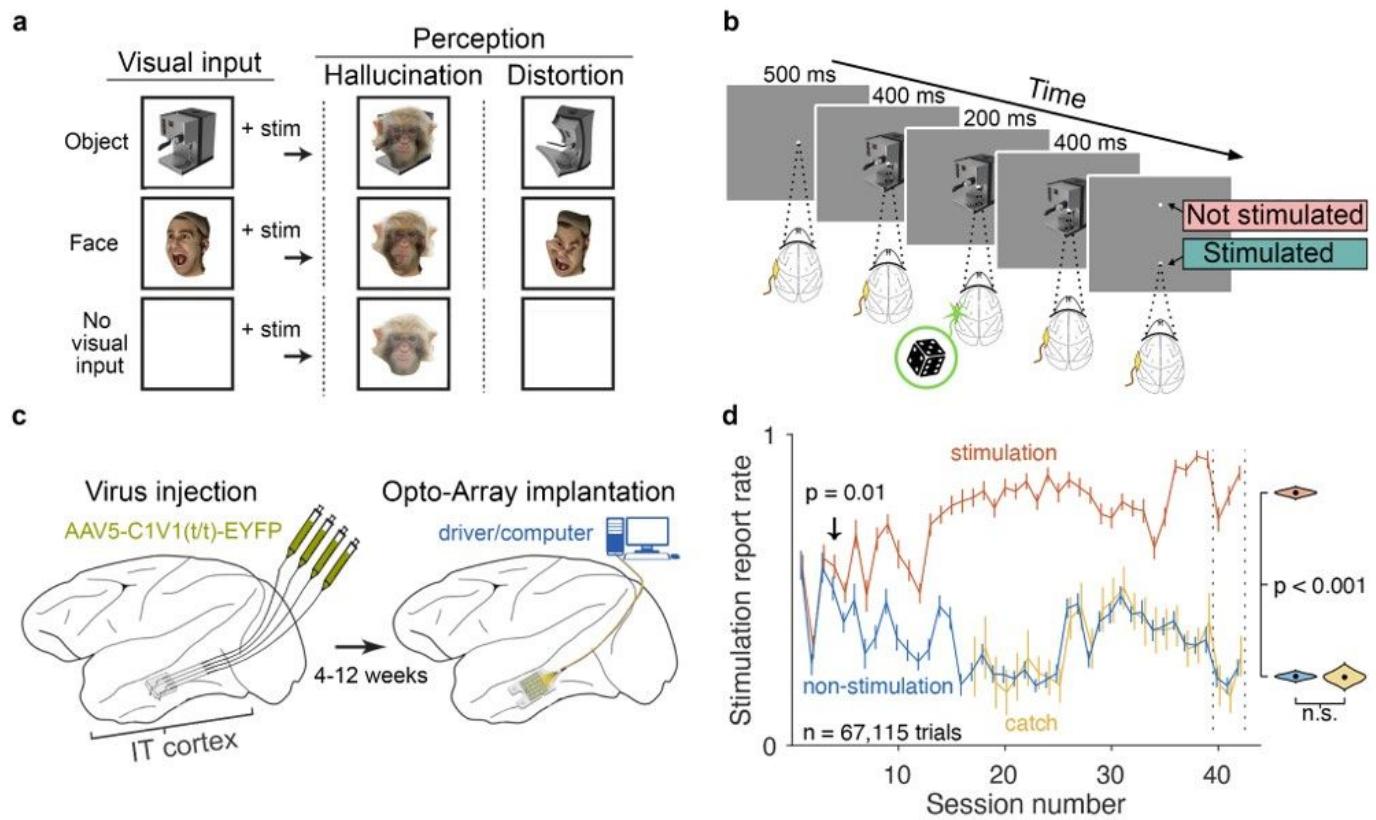


Figure 1

Hypothesis, behavioral task, surgical procedure and training phase results.

a, Schematic illustration of two hypothetical perceptual events evoked by cortical stimulation. The left column illustrates three different examples of visual stimuli presented on the screen during cortical stimulation. The middle and the right columns demonstrate the two hypotheses: The hallucination hypothesis (middle column) implies that cortical stimulation adds a specific visual element (e.g. a monkey face) to the contents of visual perception independent of the external visual stimulus. Alternatively, the distortion hypothesis (right column) assumes that stimulation-induced perceptual events highly depend on the visual input, predicting a unique event for each visual stimulus viewed. **b**, Behavioral task: in each behavioral trial following fixation an image was displayed on the screen for 1 s. In half of the trials, randomly selected, a 200 ms illumination impulse was delivered to IT cortex halfway through image presentation. The animal was rewarded for correctly identifying whether the trial did or did not contain cortical stimulation by looking at one of the two subsequently presented targets at the end of trial. **c**, Schematic illustration of the procedure for chronic optogenetic stimulation of IT cortex. Left, Injection of Adeno-associated virus (serotype 5) engineered to express the excitatory opsin C1V1. Right, Opto-Array implantation: in a separate surgery, we visually confirmed the expression of the excitatory virus and implanted an Opto-Array over the expression zone. We also implanted another array in the corresponding region of the opposite hemisphere where no virus was injected (control site, not shown). **d**,

Behavioral performance as a function of session number during the training phase. The y-axis indicates the proportion of the trials reported as stimulated. Red, blue and yellow colors represent data from the stimulation, non-stimulation, and catch trials respectively. Error bars represent bootstrapped 95% confidence intervals. The difference between stimulated and nonstimulated trials became significant at session 4 (arrow) and remained so through the training. Fluctuations of performance in time represent usage of different visual stimuli and stimulation intensities throughout the training. No significant difference was found between the catch and non-stimulation trials. The violin plots on the right side illustrate the mean and bootstrapped 95% confidence interval of stimulation report rate for each trial type in the last 3 sessions, between the dashed lines .

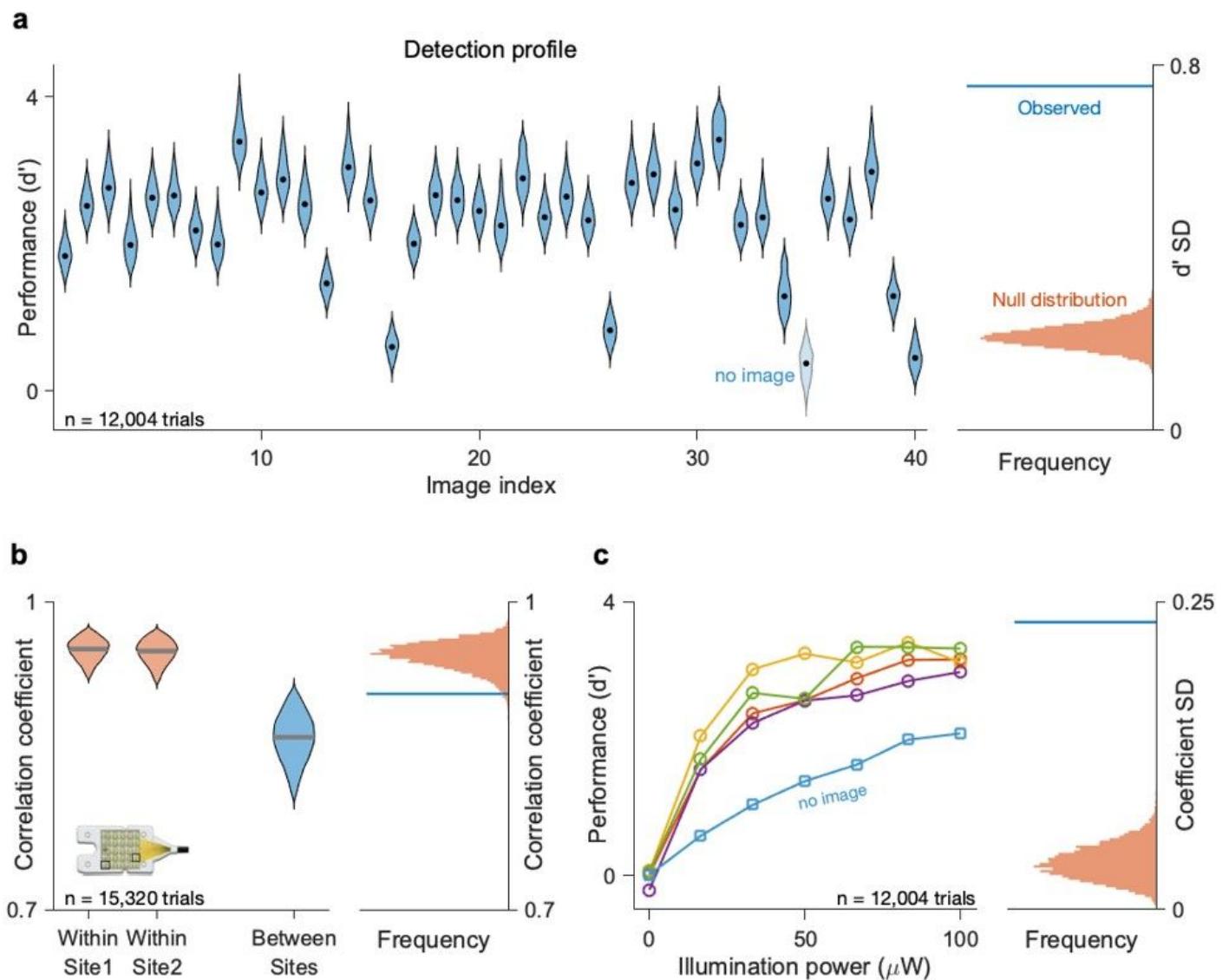


Figure 2

Stimulation detection performance is modulated by visual input, cortical location, and illumination power.

a, left, detection profile: the behavioral performance (d') on the cortical stimulation detection task for 40 images. The black dots represent d' for each image and the violin plots represent bootstrapped 95% confidence intervals. Right, permutation test: the blue line indicates the standard deviation of d' 's across images, and the red histogram represents results from a permutation test with 10,000 times randomly assigned images on trials revealing the statistical significance of the effect of image on performance. **b**, left, correlation between detection profiles within each cortical stimulation site and between them. The violin plots represent 95% confidence intervals of the bootstrapped distribution of the correlations with 10,000 resamples, and the horizontal lines indicate their medians. Right, permutation test: the blue line indicates the observed correlation between the sites. The red histogram represents results from a permutation test with 10,000 times random assignment of stimulation condition over the trials. This shows that the correlation of detection profile patterns between the sites is significantly lower than the null distribution. **c**, left, detection performance (d'), as a function of illumination power. Each line represents data from 1 image (5 images in total including the no image condition). Right, permutation test, the standard deviation of the coefficients for each image, derived from fitting of the psychometric curves. The blue line indicates the observed value, and the red distribution represents the null distribution generated by 10,000 times randomly assigning the image indexes to the trials. This confirms the coefficients are significantly different from each other.

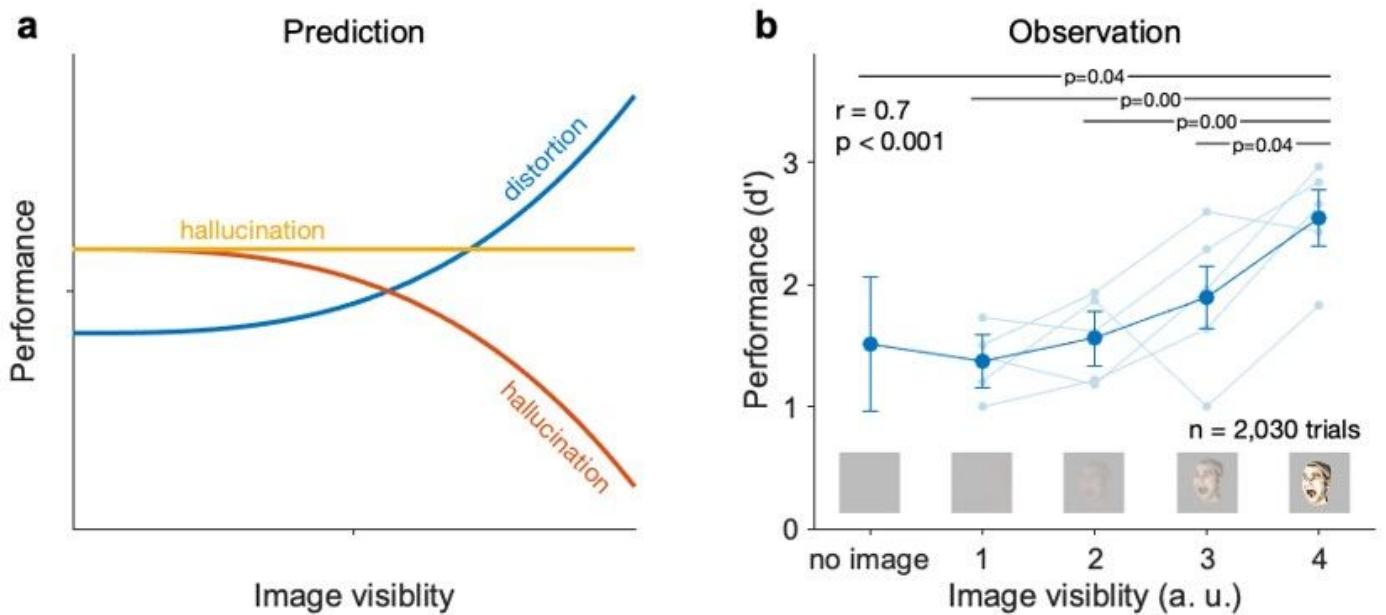


Figure 3

Stimulation detection performance is modulated by image visibility.

a, prediction: in case of hallucination, decreasing the visibility of the screen image should either not affect detectability of cortical stimulation (yellow line) or aid detection by decreasing background clutter (red line). In case of distortion, increasing the visibility should increase the detection performance (blue line),

since the perceptual effect is a function of the visual input. **b**, observation: the x-axis represents 4 levels of image visibility and the gray background, used in experiment 3. The y-axis is the detection performance (d') on the cortical stimulation detection task. The thin lines represent data from 5 different images and the thick line illustrates the overall averages. Error bars represent 95% confidence intervals. There is a significant correlation between the image visibility and performance ($r = 0.7$). The p-values for pairwise comparisons are from post-hoc tests of ANOVA (Benjamini-Hochberg corrected).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ExtendedData.docx](#)