

## Decoding perceptual awareness across the brain with a no-report fMRI masking paradigm

### Highlights

- When observers do not report their perceptual experiences, the frontal lobe is not activated
- Visibility can still be decoded in the frontal lobe even when no reports are provided
- The act of reporting versus not reporting does not affect occipitotemporal or parietal cortex

### Authors

Elaheh Hatamimajoumerd,  
N. Apurva Ratan Murty, Michael Pitts,  
Michael A. Cohen

### Correspondence

michaelthecohen@gmail.com

### In brief

Hatamimajoumerd et al. ask which parts of the brain support visual awareness. They find activation across the brain when observers report their experiences. However, without reports, the activation in the frontal lobe disappears. Meanwhile, with pattern classification, visibility can still be decoded in the frontal lobe even when no report is provided.

Article

# Decoding perceptual awareness across the brain with a no-report fMRI masking paradigm

Elaheh Hatamimajoumerd,<sup>1,2</sup> N. Apurva Ratan Murty,<sup>1,3</sup> Michael Pitts,<sup>4</sup> and Michael A. Cohen<sup>1,2,5,\*</sup>

<sup>1</sup>McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, USA

<sup>2</sup>Department of Psychology and Program in Neuroscience, Amherst College, 220 South Pleasant Street, Amherst, MA, USA

<sup>3</sup>Center for Brains, Minds, and Machines, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, USA

<sup>4</sup>Department of Psychology, Reed College, 3203 Southeast Woodstock Blvd, Portland, OR, USA

<sup>5</sup>Lead contact

\*Correspondence: [michaelthecohen@gmail.com](mailto:michaelthecohen@gmail.com)

<https://doi.org/10.1016/j.cub.2022.07.068>

## SUMMARY

Does perceptual awareness arise within the sensory regions of the brain or within higher-level regions (e.g., the frontal lobe)? To answer this question, researchers traditionally compare neural activity when observers report being aware versus being unaware of a stimulus. However, it is unclear whether the resulting activations are associated with the conscious perception of the stimulus or the post-perceptual processes associated with reporting that stimulus. To address this limitation, we used both report and no-report conditions in a visual masking paradigm while participants were scanned using functional MRI (fMRI). We found that the overall univariate response to visible stimuli in the frontal lobe was robust in the report condition but disappeared in the no-report condition. However, using multivariate patterns, we could still decode in both conditions whether a stimulus reached conscious awareness across the brain, including in the frontal lobe. These results help reconcile key discrepancies in the recent literature and provide a path forward for identifying the neural mechanisms associated with perceptual awareness.

## INTRODUCTION

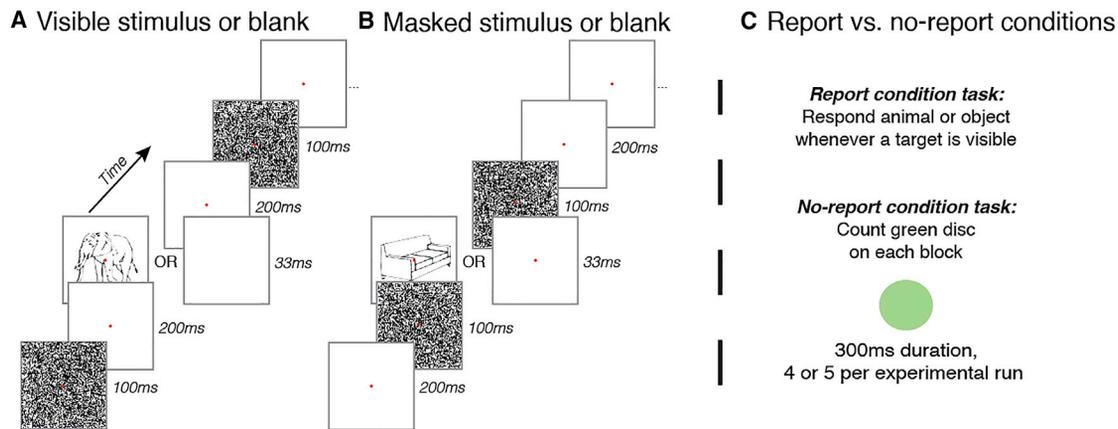
Which neural processes are critical for conscious processing in the human brain? Although this question may seem fairly simple, it has led to considerable debate over several decades. On one side of this debate are “cognitive/higher-order” theories like global neuronal workspace theory<sup>1–3</sup> or higher-order thought theory,<sup>4–6</sup> which claim that conscious experience critically depends on information reaching a fronto-parietal network, with specific emphasis placed on certain regions of the prefrontal cortex. Experimental support for this view stems from a wide variety of results using multiple paradigms showing that consciously perceived stimuli activate a fronto-parietal network, whereas unconscious stimuli only activate sensory regions (i.e., the occipital lobe and ventral temporal cortex in the case of vision).<sup>7–11</sup>

On the other side of this debate are “sensory” theories such as integrated information theory<sup>12,13</sup> or recurrent processing theory,<sup>14–18</sup> which claim that the true neural correlates of perceptual awareness are located in more posterior, sensory regions of the brain. According to these theories, activations in the frontal lobe are not associated with visual awareness but are instead associated with post-perceptual processing that occurs after an observer consciously perceives a stimulus (e.g., categorization, memory, attention, decision-making, motor outputs, etc.,<sup>16,19–21</sup>). This idea stems from the fact that in standard experimental paradigms, observers are often asked to report the contents of their perceptual experiences, which can be

accomplished in the visible conditions but not in the invisible conditions. Even in simpler detection tasks, differences in decision confidence have been shown to strongly influence differential activation in the prefrontal cortex between the seen and unseen stimuli.<sup>21</sup> Therefore, the “aware” condition is often confounded with additional steps in post-perceptual processing that the “unaware” condition is not.

This criticism led to the development of “no-report” paradigms, in which observers are aware or unaware of a stimulus, but they do not make any explicit post-perceptual judgments about that stimulus.<sup>16,21–27</sup> The goal of these no-report paradigms is to ensure that the resulting differential neural activity is exclusively associated with perceptual awareness and not with post-perceptual processing. Although there have been several studies using no-report paradigms, none so far have been used to ask the question about the role of the fronto-parietal network in conscious perception. Most of these studies either use electroencephalography (EEG), which does not have the spatial resolution to isolate specific brain regions,<sup>26,28–32</sup> or binocular rivalry, which can measure changes in the content of consciousness but cannot measure the difference between seeing a stimulus versus seeing an empty screen.<sup>33–35</sup>

To address this issue, we scanned participants with fMRI while they performed both a report and no-report task during a classic manipulation of perceptual awareness: visual masking.<sup>36,37</sup> Our primary question asked which specific neural regions differentiate between consciously and unconsciously processed visual



**Figure 1. Experimental methods**

(A) On visible trials, there were 200 ms gaps separating the stimuli from the masks.

(B) On masked trials, the masks came immediately before and after the stimulus, rendering them completely invisible.

(C) In the report condition, participants reported whether they perceived an animal or object. In the no-report condition, the stimulus presentation sequence was the same, but in this case, participants counted the number of times they saw a green disk and reported their count at the end of each block.

stimuli. The two classes of theories described above make very different predictions about what should be observed in the no-report conditions of our study. Cognitive theories predict that similar neural response patterns in fronto-parietal regions should be found in both the report and no-report conditions. Sensory theories, however, predict that posterior activations should be similar in the report and no-report conditions, whereas fronto-parietal activations should only be present in the report condition. We test these predictions using both univariate, which are the most common analysis in the literature, and multivariate analyses to investigate the neural representations supporting conscious perception in the human brain.

Overall, our results show that the differential response in the frontal lobe between visible and invisible stimuli is robust in the report condition but disappears in the no-report condition, whereas responses in parietal, ventral temporal, and occipital regions remain intact in both conditions. At first blush, this result would appear to support sensory theories over cognitive theories. Critically, however, we found that although univariate activity disappeared in the frontal lobe in the no-report condition, we could still accurately decode whether or not a stimulus was consciously perceived (i.e., visible versus invisible conditions) using multivariate analyses. Taken together, these results highlight the importance of using no-report paradigms to study the neural mechanisms of perceptual awareness, and they shed new light on the specific neural regions in both posterior and prefrontal cortices that may be critical for conscious processing, regardless of the task one is performing.

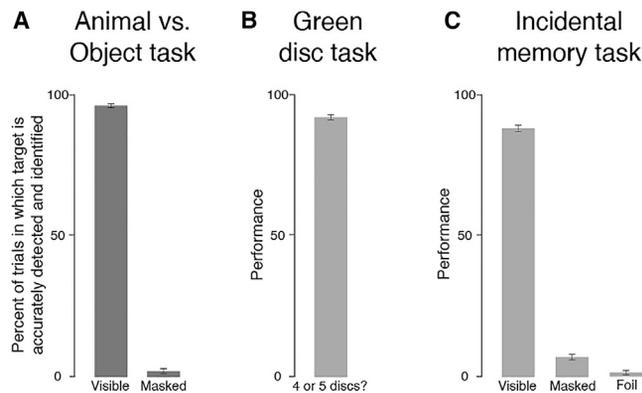
## RESULTS

### Visual masking paradigm

Our masking paradigm closely modeled the experiments of Dehaene et al.<sup>9</sup> and was previously described in an EEG experiment by Cohen et al.,<sup>30</sup> with some minor modifications. Similar to the work by Cohen et al.,<sup>30</sup> the target stimuli consisted of 32 line drawings of animals and objects (16 each) while the masks were generated using randomized line segments of the animals

and objects. Each trial consisted of a forward and backward mask, a blank gap, and either a target stimulus or a blank in between the masks (see [STAR Methods](#)). Masks were presented for 100 ms, blank gaps for 200 ms, and the critical stimuli/blanks for 33 ms. On visible trials, the 200 ms blank gaps were presented immediately before and after the presentation of the critical images (i.e., the animal and object line drawings; [Figure 1A](#)). On the invisible trials, the masks were presented immediately before and after the presentation of the critical images, rendering them invisible to observers ([Figure 1B](#)). Finally, on a subset of trials (4 or 5 times per run), a large green disk appeared in a pseudo-random fashion ([Figure 1C](#)). These green disk stimuli served as the task-relevant targets in the no-report condition and were designed to ensure spatial and temporal attention to the critical images, but they were irrelevant in the report conditions (see below). All trials containing a green disk stimulus were excluded from analyses.

Each participant performed three main tasks in the scanner: (1) report runs, (2) no-report runs, and (3) an incidental memory task. First, for the report runs, participants reported whenever they saw an animal or object at the end of each trial. On trials in which no targets were seen, participants made no response. The decision to have participants provide no response when they did not see a target was made in order to more closely match the experimental procedures previously used by Dehaene et al.<sup>9</sup> Second, for the no-report runs, participants counted the number of times the green disk appeared throughout the run and provided a response at the end of each run (either 4 or 5). It is worth stressing that all stimulus parameters were identical between the report and no-report runs, with the key difference simply being the task the participants completed (i.e., report the target category in the report runs versus count the green disks in the no-report runs). Third, for the incidental memory task, participants were shown line drawings of animals and objects and classified whether those particular drawings had been seen in the previous set of runs. It is worth emphasizing that observers were not told in advance about the memory test and did not know they would be later asked to recall what they had previously seen. During



### Figure 2. Behavioral results

In all plots, percent correct (i.e., performance) is plotted on the y axis. Error bars represent SEM.

(A) Performance on the animal/object task in the report condition. On the x axis are the different experimental conditions corresponding to when the target stimulus was visible or invisible (i.e., masked).

(B) Performance on the green-disk-counting task in the no-report condition.

(C) Performance on the incidental memory test in the no-report condition for the stimuli that were visible or masked.

these memory tests, 1/3 of the test images had been presented in the “visible” condition (i.e., lightly masked), 1/3 of the test images had been presented in the “invisible” condition (i.e., heavily masked), and 1/3 of the images had never been previously shown (i.e., foil images). Images were presented in a random order and were shown on the screen for an unrestricted duration until the participant provided a response.

Each participant completed the experiment in the same order: (1) 9 no-report runs, (2) the incidental memory test, and (3) 13 report runs, 4 of which we slightly modified and used to independently localize specific neural regions of interest<sup>38</sup> (ROIs) (see [STAR Methods](#) and below). Participants always performed the no-report runs first to minimize the possibility that participants would spontaneously categorize the targets. In other words, if the report runs came first, participants may have grown accustomed to categorizing targets as animals or objects and might carry that habit into the no-report runs, even when the targets were no longer task relevant. Although we cannot definitively conclude that participants did not do this anyway, spontaneously (i.e., the bored monkey problem,<sup>39</sup> see [discussion](#)), we opted for this design strategy to minimize that possibility.

How effectively did our different masking procedures result in stimuli being consciously or unconsciously processed (i.e., visible versus invisible)? In the report condition, participants accurately detected and classified the targets as either animals or objects 96.33% (SEM = 1.48%) of the time ([Figure 2A](#)). Conversely, participants only detected and classified the heavily masked targets 2.16% (SEM = 2.08%) of the time. In other words, stimuli in the “visible” condition were almost always perceived, whereas stimuli in the “masked” condition were almost never perceived. The fact that there was such a stark difference in performance between the “visible” and “masked” conditions in the report runs shows we can reliably render stimuli visible or invisible depending on the order and timing of the masks and blanks.

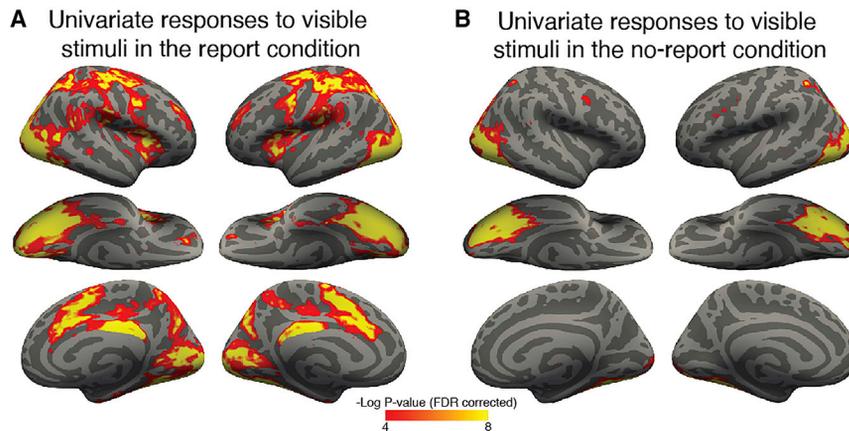
The ability to confidently control whether stimuli were visible or invisible was vital for the no-report condition, since participants never provided any reports of their perceptual experiences. Although we cannot measure how frequently participants did or did not see stimuli in the no-report runs, the results from the incidental memory task provide additional validation that observers did and did not see stimuli in the “visible” and “masked” conditions. On this task, participants correctly identified items from the “visible” condition 88% (SEM = 3.75%) of the time while only identifying items from the “invisible” condition 7.5% (SEM = 2.94%) of the time. To put that number in perspective, stimuli that had not previously been seen (i.e., foil stimuli) were claimed to have been previously seen 1.67% (SEM = 1.16%) of the time, although this number was significantly greater than the rates at which participants claimed to have “seen” masked and foil stimuli ( $t(14) = 2.83$ ,  $p = 0.01$ ).

There are two important points that should be emphasized in regard to the incidental memory task. First, these results may overestimate how often participants consciously perceive the target stimuli in the no-report runs, since each individual stimulus was presented multiple times (18 repetitions per stimulus across all no-report runs). In principle, participants could correctly identify a target during the incidental memory test, even though they only perceived that stimulus one time out of several presentations. One way to mitigate this possibility would be to increase the number of target items used and present each target only one time. However, a downside of this approach is that even if participants perceived every single unique stimulus during the experiment, they may simply forget items they saw several minutes later, especially if there were hundreds of different items. In this case, performance on the incidental memory test would incorrectly underestimate how frequently participants saw the target stimuli. Second, we did not ask participants to provide us with subjective confidence ratings in their responses to the memory tests (i.e., metacognitive reports). However, the fact that we observed a very low false alarm rate in response to the foil stimuli provides an estimate of participants’ response biases. Regardless, in the future, acquiring confidence ratings on such memory tests and exploring the trade-offs between a large stimulus set with single presentations and a small stimulus set with repeated presentations could be useful in understanding participants’ perceptual experiences in no-report situations.

### Univariate region-of-interest (ROI) analyses

Which parts of the brain are directly linked with conscious processing? A foundational result that led many researchers to believe that a fronto-parietal network plays a critical role in perceptual awareness is that these higher-level regions responded more to visible than invisible stimuli.<sup>9</sup> Thus, for our first set of analyses, we asked two basic questions: (1) can we replicate this pattern of results when observers report their perceptual experiences on a trial-by-trial basis, and (2) does this same pattern hold in a no-report condition?

To answer these questions, we first performed group-level statistics in both the report and no-report condition (see [STAR Methods](#)) In the report condition, we found that visible stimuli activated a broad fronto-parietal network, as well as the ventral occipitotemporal cortex ([Figure 3A](#)). However, the activation in the frontal lobe effectively disappeared when observers did not



**Figure 3. Neural results**

(A and B) Group random-effects maps for univariate responses to visible stimuli in both the (A) report and (B) no-report conditions across all participants (i.e., group level). On the surface, we plot all vertices for which the response was significantly greater for visible stimuli than the associated blank trials (see Figure 1A) across participants (FDR < 0.001).

have to report their experiences (Figure 3B; see Figures S1 and S2 for activation maps of every individual participant for both report and no-report conditions).

To quantify these effects, we first identified a series of ROIs in the frontal, parietal, occipital, and ventral temporal lobes. We defined these regions by contrasting neural responses when participants were shown a visible target (i.e., lightly masked) to when they were shown no target at all (i.e., a blank trial with masks only, see Figure 1A and STAR Methods). This contrast allowed us to identify voxels that responded more when participants saw a target compared with when they saw nothing at all. After identifying these voxels, they were grouped into frontal, parietal, occipital, and ventral temporal ROIs (see Figure 4A and STAR Methods for description of how voxels were grouped together). Once these regions were defined, we measured the neural responses to visible and masked stimuli after subtracting out the blank mask conditions, respectively (see STAR Methods). It should be noted that the experimental runs used to identify the ROIs were different from the runs used to then quantify the neural responses in those ROIs. Thus, all our ROIs were defined independently, and we avoided all issues of double dipping.<sup>37</sup>

Overall, we find that consciously perceived stimuli activate the occipitotemporal cortex, as well as a broad fronto-parietal network, which is consistent with prior work in which participants were tasked with reporting their experiences on a trial-by-trial basis.<sup>7–11</sup> In the report condition, we found that the neural responses elicited by visible stimuli were greater than baseline ( $t(19) > 7.13$ ,  $p < 0.001$ ) while the responses elicited by invisible stimuli were not in every ROI ( $t(19) < 2.05$ ,  $p > 0.05$ ; Figure 4B; see Table S1 for full statistics). Moreover, the response to visible stimuli was significantly greater than invisible stimuli in every ROI ( $t(19) > 6.31$ ,  $p < 0.001$  in all cases).

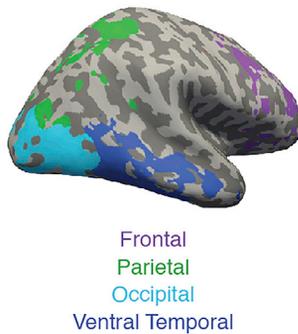
In the no-report condition, meanwhile, we observed a different pattern of results. Most importantly, we found that both visible and invisible stimuli did not significantly activate the frontal lobe more than baseline ( $t(19) > 1.11$ ,  $p > 0.28$  in both cases), and no significant difference was found between visible and invisible stimuli in this frontal region ( $t(19) = 0.39$ ,  $p = 0.35$ ; see Table S1 for full statistics); also, there was a significant interaction between visible and invisible stimuli across report and no-report conditions ( $F(1,76) = 21.82$ ,  $p < 0.001$ ). These particular results suggest that previous claims relating to the role of the

frontal lobe in conscious perception may have been somewhat premature. Indeed, our results suggest that effectively all univariate responses observed in the frontal lobe in response to visible stimuli are related to post-perceptual processing rather than conscious awareness per se. Meanwhile, in the parietal lobe, a significant response to visible stimuli was still found in the no-report condition ( $t(19) = 5.90$ ,  $p < 0.001$ ), whereas no such response was found in response to invisible stimuli ( $t(19) = 0.91$ ,  $p = 0.37$ ). Moreover, visible stimuli elicited a higher response than invisible stimuli ( $t(19) = 5.07$ ,  $p < 0.001$ ), although the size of this effect was significantly smaller in the no-report condition than in the report condition ( $F(1,76) = 17.52$ ,  $p < 0.001$ ). Finally, a similar pattern of results was seen in the no-report condition in both the occipital and ventral temporal cortices: visible stimuli activated these regions above baseline ( $t(19) > 8.85$ ,  $p < 0.001$  in both cases), invisible stimuli did not activate these regions above baseline ( $t(19) < 0.58$ ,  $p > 0.37$  in both cases), and visible stimuli elicited a higher neural response than invisible stimuli ( $t(19) > 8.78$ ,  $p < 0.001$ ).

It should be noted that these ROIs are rather large, and each occupies vast swaths of cortex. Does this pattern of results hold if we break up the cortex into smaller regions? To answer this question, we performed an exploratory analysis in which we grouped all of the significant voxels from the localizer runs, based on their location within every parcel in the Desikan-Killiany atlas ( $n = 34$ ). We then performed the same univariate analyses in each of these individual parcels. The results for each of these regions are reported in Figure S3A.

At first blush, this pattern of results appears to be consistent with predictions made by sensory theories such as the integrated information theory or recurrent processing theory, while challenging “cognitive” theories such as the global neuronal workspace theory or higher-order thought theory. In fact, the recurrent processing theory explicitly predicts that the ventral occipitotemporal cortex should differentiate between visible and masked stimuli in no-report conditions,<sup>14–18</sup> while the integrated information theory predicts that the parietal lobe would respond similarly in a no-report paradigm,<sup>12,13</sup> all of which we found in our results. However, this data alone does not “confirm” these theories or “disprove” the global neuronal workspace or higher-order thought theories. It is possible that the role of the frontal lobe in conscious processing in no-report conditions cannot be adequately captured using simple univariate measures. Indeed, the functional organization of the prefrontal cortex is likely to be somewhat random, relative to the highly organized mini-columns in posterior sensory areas.<sup>40,41</sup> Thus, it is critical to

**A** Example ROIs for one participant in report condition



examine neural responses in the frontal lobe in a no-report condition using more subtle multivariate analyses.

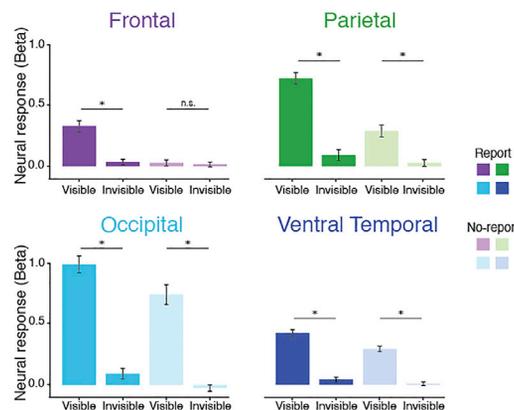
**Multivariate ROI analyses**

How well can we decode whether or not a stimulus was consciously seen (i.e., visible versus invisible stimuli) in both report and no-report conditions? To answer this question, we trained a linear support vector machine (SVM) to perform the binary classification between visible and invisible stimuli in the exact same ROIs as the univariate analyses described previously. For each run, we trained this classifier on eight experiment runs and then tested it on one held-out run not used in the training set. We repeated this analysis nine times for each individual run and averaged across runs to obtain a decoding accuracy value for every ROI in each participant for both report and no-report runs (see STAR Methods).

In the report condition, we found high decoding accuracies across every sized ROI in the frontal, parietal, occipital, and ventral temporal lobes ( $t(19) > 18.62$ ,  $p < 0.001$  in all cases; Figure 5). This finding is broadly consistent with previous work showing a significant difference in neural responses between visible and invisible stimuli across much of the cerebral cortex.<sup>1</sup> Similarly, in the parietal, occipital, and ventral temporal lobes, decoding accuracy remained above baseline in the no-report condition ( $t(19) > 19.48$ ,  $p < 0.001$  in all cases). Does this pattern of results hold in the no-report condition? Overall, we found that we could still significantly decode whether or not a stimulus was consciously processed in the frontal lobe in the no-report condition ( $t(19) = 3.90$ ,  $p < 0.001$ ). However, although decoding accuracy was significantly above chance in the no-report condition, it was also significantly lower in the no-report condition relative to the report condition ( $t(19) = 5.88$ ,  $p < 0.001$ ). A closer examination of these decoding results revealed that the classifier was able to identify those voxels that responded more to visible stimuli than invisible stimuli, as well as those that responded more to invisible stimuli than visible stimuli. In the univariate analyses, meanwhile, those differences cancelled out one another, leading to no significant overall activation in response to visible stimuli in the no-report condition.

Once again, since these ROIs are rather large and each occupies vast swaths of cortex, we repeated this decoding analysis

**B** Univariate responses in report and no-report conditions



**Figure 4. Region of interest results**

(A) Visualization of the regions of interest used for all univariate analyses in one representative subject.

(B) Univariate responses to visible and invisible/masked stimuli in both report (saturated colors) and no-report (desaturated colors) conditions for all subjects. Neural responses (i.e., beta values) are plotted on the vertical axis. Error bars represent SEM.

separately within every parcel in the Desikan-Killiany atlas ( $n = 34$ ). The results for each of these regions are reported in Figure S3B.

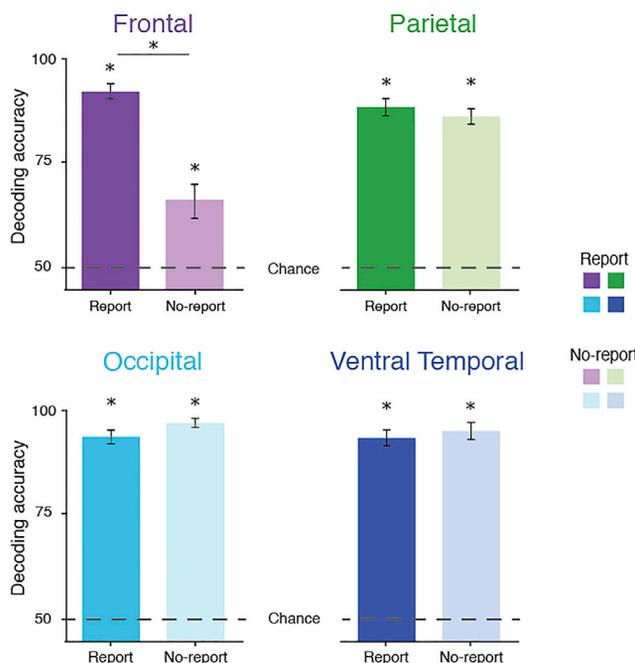
There are a few critical aspects of these results that are worth describing. First

and foremost, it is worth stressing how we could consistently decode whether or not a stimulus reached conscious awareness in the frontal lobe in the no-report condition. While this result would not naturally be predicted by sensory theories (e.g., integrated information theory, recurrent processing theory, etc.), it is a clear prediction of cognitive theories (e.g., global neuronal workspace theory, higher-order thought theory, etc.) In other words, these decoding analyses suggest that the frontal lobe might play an important role in conscious processing under conditions of no report, even though we found no univariate frontal activation in response to visible stimuli in the no-report condition (see Figures 3B and 4B). Second, although we can decode consciousness in the frontal lobe, it is the only region where decoding accuracy was consistently modulated by report versus no report. This finding is consistent with our univariate analyses, which showed drastic changes in frontal activation as a function of reporting (see Figures 3B and 4B). Thus, across both the multivariate and univariate measures, we find strong evidence that neural activity in the frontal lobe is radically altered by the act of reporting perceptual experiences. Seeing such modulation even with multivariate analyses further bolsters the idea that prior claims regarding the frontal lobe's role in conscious awareness may have been somewhat overstated, since the substantial differences between report and no-report conditions were often not considered.

**Decoding the contents of consciousness across the report and no-report conditions**

To what extent can the specific contents of perceptual experience (i.e., animals versus objects) be decoded across the brain in the report and no-report conditions? To answer this question, we trained the same linear classifier as described above in the same ROIs used in Figures 4 and 5 and asked how well the classifier could differentiate trials in which participants saw an animal from trials in which participants saw an object. Overall, we found that animals versus objects could be decoded well above chance in each of our four main ROIs in the report condition (accuracy > 86%,  $t(19) > 13.43$ ,  $p < 0.001$  in all cases; Figure S4A). Meanwhile, in the no-report condition, animals versus objects could be decoded significantly above chance in the parietal, occipital, and ventral temporal lobes (accuracy > 80%,  $t(19) > 7.82$ ,

### Multivariate decoding analyses (visible vs. masked trials)



**Figure 5. Decoding results**

Multivariate decoding analyses in both the report (saturated colors) and no-report (desaturated colors) conditions across the same regions of interest used above (Figure 4A). Decoding accuracy (visible versus masked; 50% chance) is plotted on the vertical axis. Error bars represent SEM.

$p < 0.001$  in all cases), but they could only be marginally decoded in the frontal lobe (accuracy = 58%,  $t(19) = 1.89$ ,  $p = 0.07$ ). In addition, the frontal lobe is the only region in which decoding is significantly lower in the no-report condition relative to the report condition ( $t(19) = 6.96$ ,  $p < 0.001$ ). Even though we cannot say so definitively, we believe it is likely that with more experimental runs in the no-report condition, we would be able to decode the contents of experience in the frontal lobe for the visible trials.

#### Decoding consciousness across the report and no-report conditions

Are there neural patterns associated with conscious processing that generalize across the report and no-report conditions? To answer this question, we trained the same linear classifier as described above in the same ROIs used in Figures 4 and 5 and asked how well the patterns from one condition (i.e., report) could classify patterns from the other condition (i.e., no report) and vice versa (see STAR Methods). Overall, we found that decoding consciousness could not generalize across these conditions (i.e., cross-task decoding) in the frontal lobe (train on report = 55.3% correct,  $t(19) = 1.96$ ,  $p = 0.07$ ; train on no report = 55.3%,  $t(19) = 1.40$ ,  $p = 0.17$ ; Figure S4B). However, decoding consciousness did generalize across these condition in the parietal, occipital, and ventral temporal ROIs (train on report > 73% correct,  $t(19) > 8.63$ ,  $p < 0.001$ ; train on no report = 78% accuracy,  $t(19) > 9.43$ ,  $p < 0.001$  in all cases). At first blush, it may seem

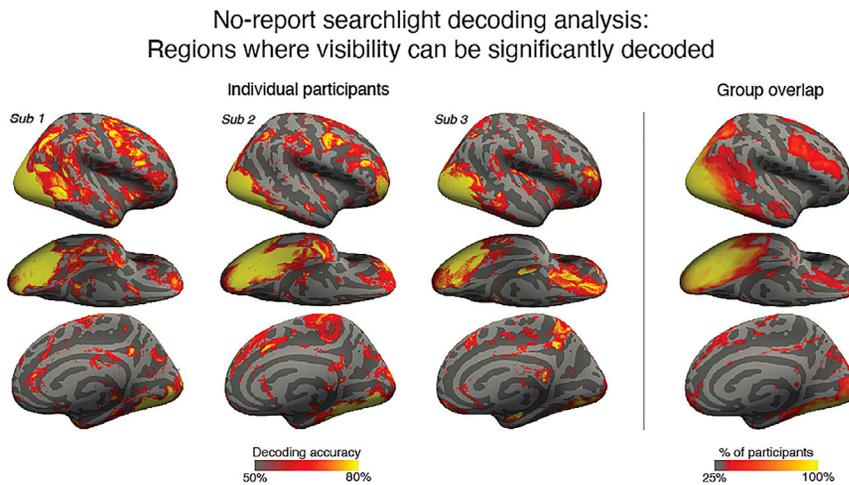
as if these results suggest that the only true neural correlates of perceptual awareness that generalize across task conditions are in more posterior regions of the brain. However, we strongly encourage readers to interpret this with extreme caution, and no conclusions are reached prematurely. It is quite likely that the reason there is no significant cross-task decoding in the frontal lobe is simply because the patterns in that specific ROI are far more contaminated by the act of report versus not reporting than in other parts of the brain. For example, in the occipital lobe, we find much less of a change in neural responses as a function of the report/no-report distinction. Therefore, it makes sense that patterns would generalize across those two conditions. In the frontal lobe, however, the patterns drastically change as a function of report/no report, making it harder for any pattern to generalize across the conditions. Thus, we strongly encourage caution when interpreting these results and are reluctant to make any definitive statements about shared patterns associated with conscious processing being found across these conditions.

#### Multivariate searchlight analyses in the no-report condition

In light of these decoding results, it is natural to wonder which specific parts of the brain, specifically the frontal lobe, are most directly involved in conscious processing in the no-report condition (i.e., dorsolateral prefrontal, medial frontal, etc.). Unfortunately, the decoding analyses described above are not well suited to answer this question. All of those analyses are carried out in large, non-contiguous ROIs comprising multiple regions scattered across each lobe of the brain (see Figure 3A). Therefore, in order to gain insight into which particular regions of the frontal lobe best differentiate between seen and unseen stimuli, we performed similar decoding analyses using a spatial searchlight approach.<sup>42</sup>

We defined a three-dimensional spherical volume around each vertex in the brain and jointly analyzed all of the vertices within that sphere (see STAR Methods). Specifically, we once again trained a linear SVM to differentiate between visible and masked trials and visualized those results across the surface of the brain. To visualize the results, we plotted the accuracy values on each vertex around which each sphere was centered. In this case, we only plotted accuracy values that were statistically above chance after using the Bonferroni correction to correct for multiple comparisons<sup>43</sup> (see STAR Methods).

Consistent with our decoding results in independently defined ROIs, we once again found significant decoding accuracies across the frontal lobe, as well as in the parietal and ventral occipitotemporal cortices (Figure 6; see STAR Methods). Critically, however, although there were regions in which conscious processing could be decoded in the frontal lobe in virtually every participant, those regions varied in their location between participants (see Figure S5 for individual participants). In spite of this variability, is it possible to identify any consistent subregions showing above-chance decoding in the frontal lobe in no-report conditions? To answer this question, we first identified every vertex for which we could decode conscious perception significantly above chance. After identifying these vertices, we overlaid these brains on top of one another to create a group overlap map. This map revealed that there was a modest amount of overlap across individuals in a dorsolateral region of the prefrontal



**Figure 6. Searchlight decoding results**

Multivariate decoding analyses using a spatial searchlight. On the left panel, regions in which visibility (i.e., visible versus invisible trials) can be significantly decoded above chance within four individual participants. On the right panel, group overlap maps where the color scale reflects the percentage of subjects for which visibility can be significantly decoded in a given location.

### Adjudicating between different theories of consciousness

At first blush, two main classes of theories relating to conscious processing in the brain make radically different predictions about the expected results for the current set of experiments. Sensory theories predict that conscious visual processing should only be associated with posterior regions in the “back of the head,”<sup>20</sup> whereas post-perceptual processing should be linked with frontal regions.<sup>12,13,21</sup> Conversely, cognitive theories predict that similar neural responses in some parts of the frontal-parietal network should still be associated with conscious processing in the no-report conditions.<sup>1,2,4–6</sup> Initially, the lack of significant univariate activation in response to visible stimuli in the no-report condition appears to be consistent with sensory theories over cognitive theories. However, the fact that we can decode conscious awareness, with multivariate analyses, in a subset of frontal regions appears more consistent with cognitive theories over sensory theories.

cortex, which is somewhat consistent with previous work associating specific prefrontal regions (e.g., in the middle and inferior frontal gyrus) with conscious processing.<sup>1,25,33,35</sup> However, it should be stressed that the overlap across participants is relatively minor and that there is significant spatial heterogeneity between individuals. It is crucial that this fact be kept in mind for future research, since group-level contrasts and analyses would likely cause researchers to incorrectly conclude that it is not possible to decode conscious perception in the frontal lobe in no-report paradigms.<sup>44,45</sup> For example, group-level statistics suggest that visibility cannot be decoded in the frontal lobe (Figure S6). This analysis on its own would lead researchers to mistakenly believe consciousness cannot be decoded in the frontal lobe in a no-report condition. In fact, consciousness can be decoded in the frontal lobe under conditions of no report so long as inter-subject variability is properly considered.

## DISCUSSION

In this experiment, we examined the response to visible and invisible stimuli in a visual masking paradigm<sup>9,30</sup> under both report and no-report conditions. Our goal was to differentiate between neural regions associated with perceptual awareness and neural regions associated with post-perceptual processing. Using basic univariate analyses, we replicated prior work showing that when observers report their experiences, neural activation extends across sensory regions and into a large-scale fronto-parietal network<sup>7–11</sup> (Figure 3A). In the no-report condition, meanwhile, differential frontal activations for visible versus invisible stimuli disappeared, and significant responses to visible stimuli were only found in parietal and ventral occipitotemporal regions (Figure 3B). However, when we asked how well we could decode visible versus invisible trials using standard multivariate analyses, we found a different pattern of results. Specifically, we found that we could still decode whether or not a stimulus was consciously perceived in the frontal lobe even under conditions of no report (although it should be noted that there was a significant decrease in decoding accuracy in the frontal lobe in the no-report condition relative to the report condition). Together, these results have significant implications for several theories of consciousness, which we discuss in detail below.

In light of these results, a natural question is, do the results presented here definitively support or challenge either class of theories? We argue that they do not and that both classes of theories can accommodate these results into their theoretical frameworks. For example, cognitive theories can easily account for the lack of neural response in the no-report paradigm using univariate measurements (Figures 3 and 4) by claiming that such measurements are simply not sensitive enough to highlight the critical role of the frontal lobe in conscious processing. Indeed, several prior studies have demonstrated how multivariate analyses can often detect differences in neural responses that univariate analyses cannot.<sup>45–48</sup> Thus, the lack of response, using univariate measurements, can still be consistent with the predictions made by cognitive theories, which would suggest that the multivariate results reflect the importance of the frontal lobe in perceptual awareness.

Conversely, sensory theories could easily account for above-chance decoding of conscious experience in the frontal lobe by appealing to “the bored monkey problem” of no-report paradigms.<sup>39</sup> This problem refers to the fact that since the task demands of the no-report runs are fairly modest (i.e., simply counting the number of green disks that appear), participants may naturally and spontaneously engage in post-perceptual cognitive processing. With minimal task demands and a fairly long experiment (i.e., 9 experimental no-report runs), participants might spontaneously think about the target animals and objects, remember seeing some of them on previous runs, muse about the contents of their experiences, and carry out other forms of

post-perceptual cognitive processing. This spontaneous cognitive processing about the stimuli would only occur in the visible condition, leading to differential neural responses between visible and invisible stimuli. Therefore, according to this line of reasoning, decoding consciousness in the frontal lobe in a no-report paradigm may not reflect perceptual awareness at all. In fact, the possibility of spontaneous post-perceptual processing has led some researchers to argue for the development of “no-cognition” paradigms that truly eliminate any possibility of post-perceptual processing.<sup>34,39</sup> Unfortunately, it is not clear how to create no-cognition versions of tasks that render stimuli either consciously or unconsciously processed (i.e., visual masking, attentional blink, change blindness, etc.). The only method for which a no-cognition version exists is binocular rivalry, which is somewhat different from these other paradigms since it is manipulating the *contents* of consciousness (subjects either see the left or right eye stimulus) rather than manipulating the *presence* of a conscious experience (subjects either see the stimulus or not). However, as there are ways in which both classes of theories can accommodate the current results, future paradigms may need to be developed to ultimately adjudicate between these diverging views of conscious processing in the brain.

### Dividing the fronto-parietal network

Generally speaking, researchers have often treated the fronto-parietal network as a single entity, and little work has been done to differentiate between the anterior and posterior regions within this network. Grouping these regions together has largely made intuitive sense given that it has been repeatedly shown that both these regions respond when a stimulus reaches conscious awareness.<sup>9,7</sup> Here, however, we report two key pieces of evidence suggesting a strong dissociation between the frontal and parietal lobes. First, using univariate analyses, we found that the parietal lobe responds to visible stimuli and distinguishes between visible and invisible stimuli in the no-report condition, whereas the frontal lobe does not (see [Figure 4B](#)). Second, while there was a significant decrease in decoding accuracy in the frontal lobe between report and no-report conditions (see [Figure 5](#)), there was effectively no such decrease in the parietal lobe. These results show how the act of reporting a perceptual experience has radically different effects on the frontal and parietal lobes, with the reporting having a substantially greater effect on the frontal lobe. More broadly, these results also indicate that researchers should exercise caution in thinking about the fronto-parietal network as a homogeneous region, and future studies should more closely examine the respective roles these regions play in conscious processing.

The potential critical role of the parietal lobe has been previously suggested by researchers who maintain that many conscious contents are specified within a posterior cortical “hot zone” comprising a temporo-parietal-occipital zone within the posterior cerebral cortex.<sup>12,20</sup> Under this view, the parietal lobe being unaffected by the report/no-report distinction is to be expected, since it is further claimed that the frontal lobe is involved in attentional control, task execution, and other forms of pre- and post-perceptual processing. Of course, it may also be the case that some of the parietal regions that are impervious to the report/no-report manipulations are also involved in higher-order processing. Indeed, it is

well established that there are regions of the parietal lobe directly involved in processes such as attentional selection,<sup>49–51</sup> working memory,<sup>52–54</sup> decision-making,<sup>55–57</sup> and perhaps modeling one’s own attention.<sup>58</sup> Therefore, the fact that the parietal lobe may be critically involved in conscious processing does not mean that consciousness can be cleanly separated from cognitive function in the brain.<sup>2,59</sup> In the future, researchers will need to examine this possibility by determining how much of this awareness-related posterior hot zone overlaps with the neural regions supporting higher-level functions, such as attention, memory, and decision-making, in the parietal lobe.

### The importance of multivariate methods

One of the key aspects of this study is the marked difference between univariate and multivariate analyses. We argue that it is imperative for researchers to consider the importance of these different analysis methods when studying the neural basis of consciousness in the future. This difference has not always been fully appreciated and has resulted in experiments that differ in whether they found a critical role of the frontal lobe in conscious processing. For example, in a series of binocular rivalry studies, Brascamp et al.<sup>34</sup> and Frässle et al.<sup>24</sup> used univariate analyses and found negligible activity in the frontal lobe associated with transitions of perception in a no-report setting. However, a series of follow-up studies in monkeys, using more sensitive multivariate analyses, found that the contents of conscious awareness could be successfully decoded in a no-report paradigm.<sup>34,60,61</sup> Of course, it is possible that the difference in conclusions from these studies has more to do with the types of neural recordings (i.e., human neuroimaging versus monkey electrophysiology) than the analysis methods. Regardless, this serves as another example, along with the present results, as to how different methodological approaches can lead to different conclusions and should be carefully assessed when interpreting overall patterns of results.

In addition to being important for future research, consideration of multivariate methods is also critical in assessing previous no-report studies. Although a wide variety of prior results have shown the importance of no-report paradigms for consciousness studies, virtually all of them fail to use more nuanced decoding measures like the ones described here,<sup>24,26–30,32,34,62</sup> but see Kapoor et al.<sup>35</sup> and Dwarakanath et al.<sup>62</sup> In fact, in a prior study from our laboratories, we found that one particular candidate signature of conscious processing, the P3b, completely disappeared in a no-report paradigm like the one described here, which goes against key predictions made by the global neuronal workspace theory.<sup>30</sup> This previous result is perfectly consistent with our univariate measures, which found that overall activity levels in the frontal lobe completely disappeared in the no-report conditions ([Figures 3 and 4](#)). Together, this collection of results suggests that the P3b may directly correspond to the univariate fMRI responses associated with reporting one’s experience. However, given the multivariate decoding analyses reported here, it is possible that applying similar analyses to EEG (or MEG) data would also find late-stage activity patterns (~300–600 ms) associated with visual consciousness (i.e., a late, “metastable” code using M/EEG).<sup>32,63–65</sup> Thus, it is entirely possible that prior data that have been used to argue against cognitive theories may yet provide evidence consistent with

the predictions made by those theories, if analyzed in a more sensitive multivariate manner.

### Future directions

We believe there are several important future directions for this work. For example, it will be important to develop no-report paradigms that render stimuli invisible via multiple mechanisms (e.g., attentional blink, change blindness, inattention blindness, interocular suppression, dichoptic fusion, etc.) in order to “triangulate” potential neural correlates of consciousness across these various paradigms.<sup>66</sup> In addition, although EEG and fMRI no-report paradigms have been previously used, the high spatial and temporal precision of intracranial recordings could provide valuable insights into the precise timing of conscious processing in the brain, including in the prefrontal cortex.<sup>35,63,67</sup> Specifically, methods with precise temporal resolution may help researchers differentiate between frontal activation associated with motor planning and execution, compared with other post-perceptual processes like decision-making, classification, memory, etc. Finally, other analysis methods, specifically representational similarity analysis, should be incorporated into no-report paradigms. Such analysis techniques could potentially allow researchers to examine the consistency of neuronal representations across report and no-report paradigms. Finally, despite some of the recent criticisms with regard to the research program aimed at identifying the neural correlates of consciousness,<sup>68,69</sup> we believe that the field is making substantial progress and that this work should continue as we steadily improve and refine our experimental paradigms and analysis methods.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Pre-registration
  - Stimuli
  - Experimental design
  - Data acquisition
  - Data preprocessing and modeling
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Defining regions of interest
  - Univariate analyses
  - Univariate group-level statistics
  - Multivariate analyses (regions of interest)
  - Multivariate analyses (cross-decoding)
  - Multivariate analyses (searchlight)

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cub.2022.07.063>.

### ACKNOWLEDGMENTS

The authors thank Nancy Kanwisher and Daniel Dennett for helpful conversations, the four anonymous reviewers for helpful feedback, and Kevin Ortego for helping develop the paradigm for an earlier EEG study. This work was supported by a National Science Foundation (Collaborative Research award no. 1829470) and a Canadian Institute for Advanced Research Fellowship to M.A.C.

### AUTHOR CONTRIBUTIONS

M.A.C. and M.P. conceptualized the study. All authors helped design the study. M.A.C. and E.H. collected the data. E.H. analyzed the data with input and supervision by M.A.C., M.P., and N.A.R.M. M.A.C. wrote the manuscript. All authors provided feedback on the final version.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 21, 2022

Revised: June 16, 2022

Accepted: July 26, 2022

Published: August 17, 2022

### REFERENCES

1. Dehaene, S., and Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227.
2. Dehaene, S. (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts* (Penguin Books).
3. Mashour, G.A., Roelfsema, P., Changeux, J.P., and Dehaene, S. (2020). Conscious processing and the global neuronal workspace hypothesis. *Neuron* 105, 776–798.
4. Lau, H., and Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends Cogn. Sci.* 15, 365–373.
5. Brown, R., Lau, H., and LeDoux, J.E. (2019). Understanding the higher-order approach to consciousness. *Trends Cogn. Sci.* 23, 754–768.
6. Odegaard, B., Knight, R.T., and Lau, H. (2017). Should a few null findings falsify prefrontal theories of conscious perception? *J. Neurosci.* 37, 9593–9602.
7. Sadaghiani, S., Hesselmann, G., and Kleinschmidt, A. (2009). Distributed and antagonistic contributions of ongoing activity fluctuations to auditory stimulus detection. *J. Neurosci.* 29, 13410–13417.
8. van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., and Roelfsema, P.R. (2018). The threshold for conscious report: signal loss and response bias in visual and frontal cortex. *Science* 360, 537–542.
9. Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D.L., Mangin, J.-F., Poline, J.-B., and Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nat. Neurosci.* 4, 752–758.
10. Sergent, C., Baillet, S., and Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nat. Neurosci.* 8, 1391–1400.
11. Lau, H.C., and Passingham, R.E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proc. Natl. Acad. Sci. USA* 103, 18763–18768.
12. Koch, C., Massimini, M., Boly, M., and Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nat. Rev. Neurosci.* 17, 307–321.
13. Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nat. Rev. Neurosci.* 17, 450–461.
14. Lamme, V.A. (2006). Towards a true neural stance on consciousness. *Trends Cogn. Sci.* 10, 494–501.

15. Lamme, V.A.F. (2010). How neuroscience will change our view on consciousness. *Cogn. Neurosci.* *1*, 204–220.
16. Lamme, V.A.F. (2018). Challenges for theories of consciousness: seeing or knowing, the missing ingredient and how to deal with panpsychism. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* *373*, 1755.
17. Block, N. (2005). Two neural correlates of consciousness. *Trends Cogn. Sci.* *9*, 46–52.
18. Zeki, S., and Bartels, A. (1999). Toward a theory of visual consciousness. *Consc. Cogn.* *8*, 225–259.
19. Aru, J., Bachmann, T., Singer, W., and Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neurosci. Biobehav. Rev.* *36*, 737–746.
20. Boly, M., Massimini, M., Tsuchiya, N., Postle, B.R., Koch, C., and Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. *J. Neurosci.* *37*, 9603–9613.
21. Tsuchiya, N., Wilke, M., Frässle, S., and Lamme, V.A.F. (2015). No-report paradigms: extracting the true neural correlates of consciousness. *Trends Cogn. Sci.* *19*, 757–770.
22. Mazor, M., Dijkstra, N., and Fleming, S.M. (2022). Dissociating the neural correlates of subjective visibility from those of decision confidence. *J. Neurosci.* *42*, 2562–2569.
23. Farooqui, A.A., and Manly, T. (2018). When attended and conscious perception deactivates fronto-parietal regions. *Cortex* *107*, 166–179.
24. Frässle, S., Sommer, J., Jansen, A., Naber, M., and Einhäuser, W. (2014). Binocular rivalry: frontal activity relates to introspection and action but not to perception. *J. Neurosci.* *34*, 1738–1747.
25. Wiegand, K., Heiland, S., Uhlh, C.H., Dykstra, A.R., and Gutschalk, A. (2018). Cortical networks for auditory detection with and without informational masking: task effects and implications for conscious perception. *Neuroimage* *167*, 178–190.
26. Pitts, M.A., Padwal, J., Fennelly, D., Martinez, A., and Hillyard, S.A. (2014). Gamma-band activity and the P3 reflect post-perceptual processes, not visual awareness. *Neuroimage* *101*, 337–350.
27. Pitts, M.A., Martínez, A., and Hillyard, S.A. (2012). Visual processing of contour patterns under conditions of inattention blindness. *J. Cogn. Neurosci.* *24*, 287–303.
28. Shafto, J.P., and Pitts, M.A. (2015). Neural signatures of conscious face perception in an inattention blindness paradigm. *J. Neurosci.* *35*, 10940–10948.
29. Koivisto, M., Grassini, S., Salminen-Vaparanta, N., and Revonsuo, A. (2017). Different electrophysiological correlates of visual awareness for detection and identification. *J. Cogn. Neurosci.* *29*, 1621–1631.
30. Cohen, M.A., Ortego, K., Kyroudis, A., and Pitts, M. (2020). Distinguishing the neural correlates of perceptual awareness and post-perceptual processing. *J. Neurosci.* *40*, 4925–4935.
31. Schlossmacher, I., Dellert, T., Pitts, M., Bruchmann, M., and Straube, T. (2020). Differential effects of awareness and task relevance on early and late ERPs in a no-report visual oddball paradigm. *J. Neurosci.* *40*, 2906–2913.
32. Sergent, C., Corazzol, M., Labouret, G., Stockart, F., Wexler, M., King, J.-R., Meyniel, F., and Pressnitzer, D. (2021). Bifurcation in brain dynamics reveals a signature of conscious processing independent of report. *Nat. Commun.* *12*, 1149.
33. Safavi, S., Kapoor, V., Logothetis, N.K., and Panagiotaropoulos, T.I. (2014). Is the frontal lobe involved in conscious perception? *Front. Psychol.* *5*, 1063.
34. Brascamp, J., Blake, R., and Knapen, T. (2015). Negligible fronto-parietal BOLD activity accompanying unreportable switches in bistable perception. *Nat. Neurosci.* *18*, 1672–1678.
35. Kapoor, V., Dwarakanath, A., Safavi, S., Werner, J., Besserve, M., Panagiotaropoulos, T., and Logothetis, N. (2022). Decoding the contents of consciousness from prefrontal ensembles. *Nat. Commun.* *13*, 1535.
36. Kouider, S., and Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* *362*, 857–875.
37. Breitmeyer, B.G. (2008). Visual masking: past accomplishments, present status, future developments. *Adv. Cogn. Psychol.* *3*, 9–20.
38. Vul, E., Harris, C., Winkielman, P., and Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect. Psychol. Sci.* *4*, 274–290.
39. Block, N. (2019). What is wrong with the no-report paradigm and how to fix it. *Trends Cogn. Sci.* *23*, 1003–1013.
40. Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* *24*, 167–202.
41. Miller, E.K. (2000). The prefrontal cortex and cognitive control. *Nat. Rev. Neurosci.* *1*, 59–65.
42. Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. *Proc. Natl. Acad. Sci. USA* *103*, 3863–3868.
43. Bonferroni, C.E. (1936). Teoria statistica delle classi e calcolo delle probabilità. *Pubbl. R Ist. Super. Sci. Econ. Comm. Firenze* *8*, 3–62.
44. Saxe, R., Brett, M., and Kanwisher, N. (2006). Divide and conquer: a defense of functional localizers. *Neuroimage* *30*, 1088–1096. discussion 1097.
45. Fedorenko, E., and Kanwisher, N. (2009). Neuroimaging of language: why hasn't a clear picture emerged? *Lang. Linguist. Compass* *3/4*, 839–865.
46. Norman, K.A., Polyn, S.M., Detre, G.J., and Haxby, J.V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* *10*, 424–430.
47. Sterzer, P., Haynes, J.D., and Rees, G. (2008). Fine-scale activity patterns in high-level visual areas encode the category of invisible objects. *J. Vis.* *8*, 1–12.
48. Raizada, R.D.S., and Kriegeskorte, N. (2010). Pattern-information fMRI: new questions which it opens up and challenges which face it. *Int. J. Imaging Syst. Technol.* *20*, 31–41.
49. Kanwisher, N., and Wojciulik, E. (2000). Visual attention: insights from brain imaging. *Nat. Rev. Neurosci.* *7*, 91–100.
50. Chun, M.M., Golomb, J.D., and Turk-Browne, N.B. (2011). A taxonomy of external and internal attention. *Annu. Rev. Psychol.* *62*, 73–101.
51. Corbetta, M., and Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* *3*, 201–215.
52. Todd, J.J., and Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* *428*, 751–754.
53. Xu, Y. (2017). Reevaluating the sensory account of visual working memory storage. *Trends Cogn. Sci.* *21*, 794–815.
54. Serences, J.T. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Res.* *128*, 53–67.
55. Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* *86*, 1916–1936.
56. Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* *30*, 535–574.
57. Wagner, A.D., Shannon, B.J., Kahn, I., and Buckner, R.L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* *9*, 445–453.
58. Graziano, M.S. (2019). *Consciousness and the Social Brain* (Oxford University Press).
59. Cohen, M.A., and Dennett, D.C. (2011). Consciousness cannot be separated from function. *Trends Cogn. Sci.* *15*, 358–364.
60. Panagiotaropoulos, T.I., Dwarakanath, A., and Kapoor, V. (2020). Prefrontal cortex and consciousness: beware of the signals. *Trends Cogn. Sci.* *24*, 343–344.
61. Dwarakanath, A., Kapoor, V., Werner, J., Safavi, S., Fedorov, L.A., Logothetis, N.K., and Panagiotaropoulos, T. (2020). Prefrontal state fluctuations control access to consciousness. Preprint at bioRxiv. <https://doi.org/10.1101/2020.01.29.924928>.

62. Dellert, T., Müller-Bardorff, M., Schlossmacher, I., Pitts, M., Hofmann, D., Bruchmann, M., and Straube, T. (2021). Dissociating the neural correlates of consciousness and task relevance in face perception using simultaneous EEG-fMRI. *J. Neurosci.* *41*, 7864–7875.
63. Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., Hasboun, D., Baulac, M., Cohen, L., and Naccache, L. (2009). Converging intracranial markers of conscious access. *PLoS Biol.* *7*, e61.
64. King, J.R., Pescetelli, N., and Dehaene, S. (2016). Brain mechanisms underlying the brief maintenance of seen and unseen sensory information. *Neuron* *92*, 1122–1134.
65. Marti, S., and Dehaene, S. (2017). Discrete and continuous mechanisms of temporal selection in rapid visual streams. *Nat. Commun.* *8*, 1955.
66. Morales, J., Odegaard, B., and Maniscalco, B. (2019). The neural substrates of conscious perception Without performance confounds. Preprint at PsyArxiv. <https://doi.org/10.31234/osf.io/8zhy3>.
67. Kronember, S., Aksent, M., Ding, J., Ryu, J.H., Xin, Q., Ding, Z., Prince, J., Kwon, H., Khalaf, A., Forman, S., et al. (2021). Brain networks in human conscious visual perception. Preprint at bioRxiv. <https://doi.org/10.1101/2021.10.04.462661>.
68. Rahimian, S. (2022). The myth of when and where: how false assumptions still haunt theories of consciousness. *Consc Cogn.* *97*, 103246.
69. Revach, D., and Salti, M. (2021). Expanding the discussion: revision of the fundamental assumptions framing the study of the neural correlates of consciousness. *Consc Cogn.* *96*, 103229.
70. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* *12*, 2825–2830.
71. Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., and Broussard, C. (2007). What's new in psychtoolbox-3. *Perception* *36*, 1.
72. Kay, K.N., Rokem, A., Winawer, J., Dougherty, R.F., and Wandell, B.A. (2013). GLMdenoise: a fast, automated technique for denoising task-based fMRI data. *Front. Neurosci.* *7*, 247.
73. Charest, I., Kriegeskorte, N., and Kay, K.N. (2018). GLMdenoise improves multivariate pattern analysis of fMRI data. *Neuroimage* *183*, 606–616.
74. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* *31*, 968–980.
75. Hatamimajoumerd, E., Talebpour, A., and Mohsenzadeh, Y. (2020). Enhancing multivariate pattern analysis for magnetoencephalography through relevant sensor selection. *Int. J. Imaging Syst. Technol.* *30*, 473–494.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Group-level Data	This paper	<a href="https://osf.io/8zg6b/files/osfstorage/62d9c43d1bb7a54fc11f3a4a">https://osf.io/8zg6b/files/osfstorage/62d9c43d1bb7a54fc11f3a4a</a>
Univariate response data	This paper	<a href="https://osf.io/8zg6b/files/osfstorage/62d9c46627b74610db0abfd4">https://osf.io/8zg6b/files/osfstorage/62d9c46627b74610db0abfd4</a>
Searchlight analysis data	This paper	<a href="https://osf.io/8zg6b/files/osfstorage/62d9c4551bb7a54fbe1f37e4">https://osf.io/8zg6b/files/osfstorage/62d9c4551bb7a54fbe1f37e4</a>
Multivariate analysis data	This paper	<a href="https://osf.io/8zg6b/files/osfstorage/62d9c415c79a4c450d9e5a76">https://osf.io/8zg6b/files/osfstorage/62d9c415c79a4c450d9e5a76</a>
Cross decoding data	This paper	<a href="https://osf.io/8zg6b/files/osfstorage/62d9c3ff27b74610e00ac28d">https://osf.io/8zg6b/files/osfstorage/62d9c3ff27b74610e00ac28d</a>
Animal/object decoding	This paper	<a href="https://osf.io/waph2">https://osf.io/waph2</a>
Software and algorithms		
Freesurfer	Laboratory for Computational Neuroimaging at the Martinos Center for Biomedical Imaging.	<a href="https://surfer.nmr.mgh.harvard.edu/">https://surfer.nmr.mgh.harvard.edu/</a>
MATLAB	The MathWorks, Inc.	<a href="https://www.mathworks.com/">https://www.mathworks.com/</a>
Scikit-learn (Machine Learning package in Python)	Pedregosa et al., <sup>70</sup>	<a href="https://github.com/scikit-learn/scikit-learn">https://github.com/scikit-learn/scikit-learn</a>

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Michael A. Cohen ([michaeltcohen@gmail.com](mailto:michaeltcohen@gmail.com)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

- De-identified results from the ROI analyses have been deposited at osf.io and are publicly available as of the date of publication. Accession numbers are listed in the [key resources table](#).
- All original code has been deposited at osf.io and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

Twenty subjects participated in the experiment (14 females and 6 males). Two additional subjects were recruited but had to be excluded for falling asleep in the scanner and being unable to see the visual stimuli. The experiment was approved by the Committee on the Use of Humans as Experimental Subjects of the Massachusetts Institute of Technology (MIT). Participants provided informed written consent before the experiment and were compensated for their time.

### METHOD DETAILS

#### Pre-registration

All of the experimental methods and main analyses were pre-registered and are available on the Open Science Framework: <https://osf.io/8zg6b/>

#### Stimuli

The target stimuli were 32 line drawings, 16 objects and 16 animals. Each participant was shown all 32 stimuli over the course of the experiment. The stimuli were divided into 4 groups (A, B, C, and D), each group consisting of 4 animals and 4 objects. Each participant, according to their subject number, was assigned a stimulus group (A, B, C, or D) for each combination of trial type and task

condition: (1) visible, report; (2) masked, report; (3) visible, no-report; and (4) masked, no-report. These stimulus groups were counter-balanced across participants using a Williams Latin square design (Williams, 1949). Therefore, in each task condition (report, no-report), exactly half of the stimuli were presented to each participant. Due to the manipulation of awareness via masking, exactly half of the presented stimuli were visible within each task condition.

The masks were constructed using line segments from the animal and object stimuli, overlaid upon each other such that no obvious shapes could be perceived. A total of 8 mask variants were used (created by rotating and flipping the original mask), and a pair of non-matching masks was randomly selected for each trial. All stimuli and masks were 625x625 pixel (px) images. Large green discs (RGB: 200,255,200) with a diameter of 750 px, served as the target stimuli in the no-report condition.

Stimuli were controlled using Psychophysics Toolbox Version 3 for MATLAB.<sup>71</sup> All stimuli were presented on a white background (BenQ 120Hz monitor, 1920x1080 px). Participants were seated 70 cm from the monitor, thus making the average size of the critical stimuli 8.17°, the masks 14.05°, and the green discs 16.26°. All stimuli were presented at the center of the screen, while participants were instructed to maintain fixation on a small red fixation dot (0.20°) that was continuously visible throughout the procedure. All of the MATLAB code and image files needed to run the experiment are available on the Open Science Framework: <https://osf.io/8zg6b/>

### Experimental design

Each participant completed all 4 parts of the experimental procedure in the same order: (1) 9 no-report runs, (2) an incidental memory test on those task-irrelevant stimuli, (3) 9 report runs, and (4) 4 localizer runs (22 total experimental runs).

#### No-report runs & report runs

It is worth stressing that the stimulus presentation procedures were identical between no-report and report runs. The only differences were a) the tasks participants were asked to perform (see main text) and b) the stimuli used (see [stimuli](#) section above). Each trial consisted of a series of four masks and three targets in rapid succession with blanks interspersed to differentiate visible and masked trials (see [Figure 1](#) in main text). The decision to present each target 3 times on each trial was simply to try and increase the signal strength associated with each target in our event-related design. In the report runs, observers would wait until the 3<sup>rd</sup> presentation of the target to report having seen an animal or object. In the no-report runs, no such report was ever made.

In each run, there were always either 36 or 37 total trials: 8 trials where the target was present and was visible, 8 trials where the target was present and was invisible/masked, 8 trials where the target was not present with longer inter-stimulus intervals between the masks that would correspond to visible trials, 8 trials where the target was not present with shorter inter-stimulus intervals between the masks that would correspond to invisible/masked trials, and either 4 or 5 trials in which a green disc was presented (see [Figure 1A](#) in main text). For every green disc trial, there was a ¼ possibility that it would occur with each of the different other trial types (i.e., visible, masked, blank, etc.). All trials in which a green disc appeared were excluded from all fMRI analyses in both the no-report and report runs. The order in which the trials were presented was determined by Optseq (<http://surfer.nmr.mgh.harvard.edu/optseq>). Similarly, the timing between trials was determined by Optseq with anywhere between 0 and 12 seconds in between trials (though on average there was rarely long gaps between trials).

On the report runs, participants reported if they perceived an animal or object after every trial. Responses were provided by pressing two possible buttons on a button box that was held in the right hand (i.e., one button corresponded to animals, the other button corresponded to objects). The mapping between target category and button was counterbalanced across participant but remained constant within each participant for all runs. On trials in which no animal or object was perceived, participants made no such response. No feedback was ever provided at any point during the experiment.

#### Incidental memory test

See main text for description of incidental memory test.

#### Localizer runs

The localizer runs were report runs, in which observers would report every time they saw an animal or object. However, the design of these runs different from the 9 report runs described above in a few key ways. First, the only types of trials shown in these runs were trials where the target was present and was visible, as well as trials where the target was not present but with longer inter-stimulus intervals between the masks that would correspond to visible trials (left panel of [Figure 1A](#)). Thus, when localizing regions, we were ultimately contrasting visible trials vs blank trials. Second, rather than use an event-related design, these runs employed a block design in order to increase the relative power. In each run, there were 3 visible blocks and 3 blank blocks. Moreover, within each block, there were 8 trials.

#### Data acquisition

All experiments were performed at the Athinoula A. Martinos Imaging Center at MIT on a Siemens 3-T MAGNETOM Prism Scanner with a 32-channel head coil. We acquired a high-resolution T1-weighted (multi echo MPRAGE) anatomical scan during the first scanning session (acquisition parameters: 176 slices, voxel size: 3x3x3 mm, repetition time [TR]=2,500 ms, echo time [TE]=2.9 ms, flip angle=8°). Functional scans included 120 and 140 T2\*-weighted echoplanar blood-oxygen level-dependent (BOLD) images for each experimental and localizer run, respectively (acquisition parameters: simultaneous interleaved multi-slice acquisition 2, TR=2,000 ms, TE=30 ms, voxel-size 2-mm isotropic, number of slices=52, flip angle: 90°, echo-spacing 0.54 ms, 7/8 phase partial Fourier acquisition).

### Data preprocessing and modeling

fMRI data preprocessing was performed on Freesurfer (version:6.0.0; Downloaded from: <https://surfer.nmr.mgh.harvard.edu/fswiki/>). Data pre-processing included slice time correction, motion correction of each functional run, alignment to each subject's anatomical data, and smoothing using a 5 mm FWHM Gaussian kernel. Generalized linear modelling (GLM) for the dynamic localizer was also performed on Freesurfer and included 5 regressors — 1) visible target present, 2) visible blank, 3) masked target present, 4) masked blank, and 5) and green disc trials — as well as nuisance regressors for linear drift removal and motion correction per run, and analyzed on the surface reconstructed versions of the data. GLM analysis for the event-related experiment was performed using GLMdenoise.<sup>72</sup> This method, optimized for event-related fMRI, estimates the noise regressors directly from the data. Consistent with previous reports,<sup>72,73</sup> this method substantially improved the test-reliability of the estimated beta parameters in our pilot experiments. Using this method, we estimated a single beta parameter estimate for each stimulus for each run in response to each respective stimulus condition.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Defining regions of interest

To localize the regions of interest, we first calculated the contrast map of visible stimuli versus visible blanks using localizer runs and select all voxels that responded to visible stimuli significantly more than blank stimuli ( $P < 0.05$ ). To constrain these contrasts, we used the Desikan-Killiany cortical Atlas<sup>74</sup> and merged multiple parcels within each lobe to form each ROI (Frontal lobe: medial orbitofrontal, lateral orbitofrontal, superior frontal, middle frontal, orbitalis, triangularis, opercularis, and caudal middle frontal; Parietal: inferior parietal and superior parietal; Occipital: pericalcarine and lateral occipital; Ventral Temporal: fusiform, parahippocampal, inferior temporal, and middle temporal).

### Univariate analyses

For each participant, the subject-specific beta weights for all different stimuli conditions were derived through a general linear model (GLM) for every experimental run. Betas were extracted for each run for each category (visible when the target is presented, visible blank, invisible when the target is presented and invisible blank) for both report and no-report separately. We then removed the mask effect by subtracting the beta weights of visible and invisible blank conditions from the visible and invisible when the target is presented, we averaged the resulting beta weights across the predefined voxels in all different ROIs.

### Univariate group-level statistics

To create group-level significance maps of the univariate data, we first transformed every individual participant's data to fit onto the FsAverage brain and all subsequent analyses were then conducted in this anatomical space. Then, we subtracted the neural response to blank stimuli (i.e., only masks) from the neural response to visible stimuli (i.e., lightly masked) in every individual participant. Thus, for every vertex, we have 20 values that correspond to the difference in response for every individual participant. At each vertex, we then performed a one-sided *t*-test to determine if the average response of a given vertex was significantly greater than zero. Finally, we used a False Discovery Rate (FDR) threshold of 0.001 to correct for multiple comparisons and visualized those vertices that were above this threshold on the surface (Figure 3).

### Multivariate analyses (regions of interest)

For this analysis, we examined the entire multi-dimensional patterns of beta weights and then used a classifier to try and successfully decode visible from invisible/masked trials. Given the 9 experimental runs per paradigm (i.e., report vs. no-report) and two stimulus conditions (i.e., visible vs. invisible/masked) within each run, we have 18 beta vectors corresponding to visible and invisible trials for the report runs and no-report runs respectively (36 total across both report and no-report). For each participant, we first performed a principal component analysis (PCA) to avoid the curse of dimensionality and selected the number of principal components that explained 99% of data variance<sup>75</sup> (~12-15 components). We then trained a linear SVM classifier on the resulting data through a leave-one-out cross validation procedure in which one sample was used as a test and the remaining 17 samples were used as training set. This process was then repeated 18 times to make sure all the samples were used in training and test sets respectively. To calculate classifier accuracy for each participant, we simply averaged the test accuracy across all 18 iterations.

### Multivariate analyses (cross-decoding)

For this analysis, we took the same data as used in the multivariate ROI analysis (i.e., 18 beta vectors corresponding to visible and invisible trials for the report runs and no-report runs respectively (36 total across both report and no-report) that was distilled down to ~12-15 components via PCA. The only difference for this analysis is that the held out run that would be tested (and was not part of the training batch) came from a different condition. For example, if we trained on 17 samples from the report condition, we would then test the classifier on one held out sample from the no-report condition. As described above, this process was repeated 18 times so that all the samples were used in training and test sets respectively. To calculate classifier accuracy for each participant, we simply averaged the test accuracy across all 18 iterations.

### Multivariate analyses (searchlight)

We used this classification procedure within a spatial searchlight to create maps for classifying visible stimuli from the invisible ones across the cortical surfaces for both report and no report paradigm for each participant. In this case, all individual participant brains were transformed to fit onto the FsAverage brain and all searchlight analyses were conducted in this anatomical space. First, for each vertex on the surfaces, we defined a local searchlight sphere consisting of a center vertex and its neighborhood within a 1000-vertex radius using the standard Euclidean distance. This resulted in the formation of 327,684 spheres throughout the whole brain. Similar to ROI decoding, we applied PCA dimensionality reduction on each sphere before the classification. We then fed the resulting data of each sphere to a linear SVM classifier and evaluated the classifier using a repeated hold out method in which the dataset was randomized and partitioned into two sets of train and test (13 runs were used for training and the remaining 5 were used as a test set). This process was then repeated, (i.e., random sampling and classifying) 100 times and we then averaged across those 100 values to get the average classification accuracy for each vertex.

To test the statistical significance of each classification accuracy at each location, we generated a null distribution. To generate the null distribution, we first repeated the same process of train and test on 13 runs (train) and 5 runs (test), except this time we shuffled the labels of the data. This process was also repeated 100 times and then we averaged across those 100 values to get a null value at each vertex. Then, in order to test the statistical significance of each vertex, we compared the 100 values from the decoding procedure (see paragraph above) to 100 randomly sampled null values from across the brain (~50% accuracy for each distribution). With 100 values from the original data and the null distribution, we performed a two-sample *t*-test at every vertex and used the Bonferroni correction with the threshold of  $(0.05/327,648 = \text{adjusted significance threshold of } P < 0.00000015)$ .