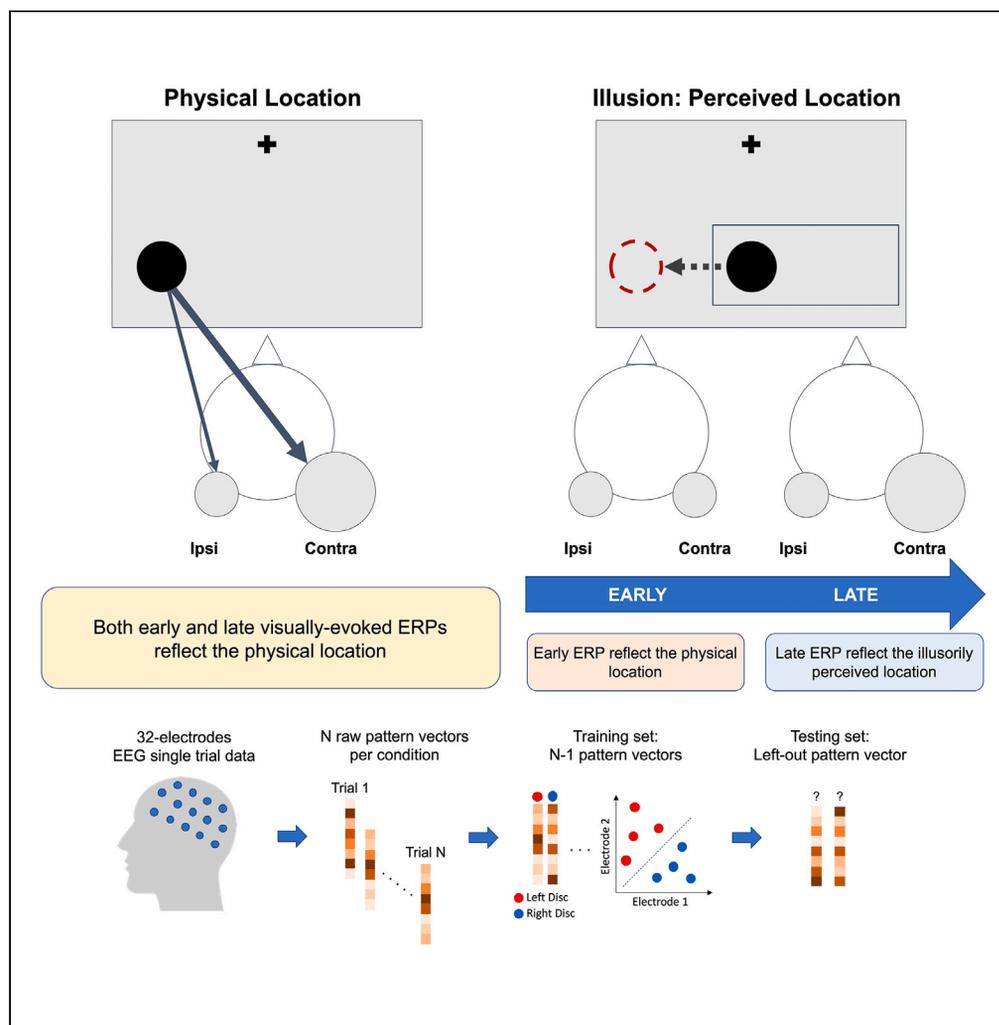**Article**

Unveiling the time course of visual stabilization through human electrophysiology

Yong Hoon Chung, Viola S. Störmer

yong.hoon.chung.gr@dartmouth.edu

Highlights

Visually evoked responses reflect illusorily stabilized positions after 140 ms

Neural codes are shared between illusory and real positions later in time

Visual stabilization involves neural feedback processes

Article

Unveiling the time course of visual stabilization through human electrophysiology

Yong Hoon Chung[1,2,*] and Viola S. Störmer[1]

SUMMARY

Object positions are coded relative to their surroundings, presumably providing visual stability during eye movements. But when does this perceived stability arise? Here we used a visual illusion, the frame-induced position shift, and measured electrophysiological activity elicited by an object whose perceived position was either shifted because of a surrounding frame or not, thus dissociating perceived and physical locations. We found that visually evoked responses were sensitive to only physical location earlier in time ($\sim$70 ms), but both physical and illusory location information was present at a later time point ($\sim$140 ms). Furthermore, location information could be reliably decoded across physical and illusory locations during the later time interval but not during the earlier time interval, demonstrating that neural activity patterns are shared between the two processes at a later stage. These results suggest that visual stability of objects emerges relatively late and is thus dependent on recurrent feedback from higher processing stages.

INTRODUCTION

In everyday life, we constantly encounter dynamic visual inputs, especially because of frequent eye movements. However, despite the constant shifts on our retinas, the world appears mostly stable to us. One way to deal with such dynamicity in visual inputs is to code the information relative to the surrounding environment as the relational information between visual stimuli is often much more stable than the location information of an individual stimulus. For instance, when looking at a book on top of a table, the relational coding of the location of the book and the table will not change even when moving your eyes and shifting the retinal locations of the two objects. This allocentric processing has been shown to influence visual perception in a variety of ways, such as perceiving the direction and velocity of object motion,[1,2] the sense of heading direction,[3] and processing of object orientation.[4]

One recent study by Özkan et al.[5] demonstrates a particularly strong example of how relational position coding can affect perception. This so-called Frame-Induced Position Shift (FIPS) illusion elicits a paradoxical stabilization of a moving object: When a frame is moved left to right and two probe stimuli are flashed inside the frame at the exact same physical location shortly before and after the frame moves, participants report these probes to appear at separate locations that match the distance that the frame traveled.[5,6] This perceived position shift occurs robustly across different visual environments such as apparent motion of a frame (instead of actual motion), whole background shifts instead of a moving frame (i.e., background pattern moving left to right), or 3D frame flipping motion.[7] In all cases, this frame-induced position shift yields a drastic perceptual effect, in some conditions resulting in 100% stabilization: The locations of dots are perceived in the frame's coordinates as if the frame was not in motion at all. Such dramatic effects of perceived position shifts were taken as evidence that surrounding context (such as a moving frame) may act to stabilize the perception of relative locations during eye movements.[5]

But how and when does this perceived stability of object locations arise? Previous research has used electroencephalography (EEG) to examine the time course of when perceived (relative to physical) location processing arises in the visual processing stream in other paradigms that induce position shifts. For example, one study used the flash-grab effect, a motion-induced position shift where moving stimuli make flashing probes appear at projected locations, and showed that illusory positions can be successfully decoded as early as 81 ms after stimulus.[8] Other investigations on the neural processes of this illusory position shift have used functional Magnetic Resonance Imaging (fMRI) and multivariate pattern analyses

[1]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA

[2]Lead contact

*Correspondence: yong.hoon.chung.gr@dartmouth.edu

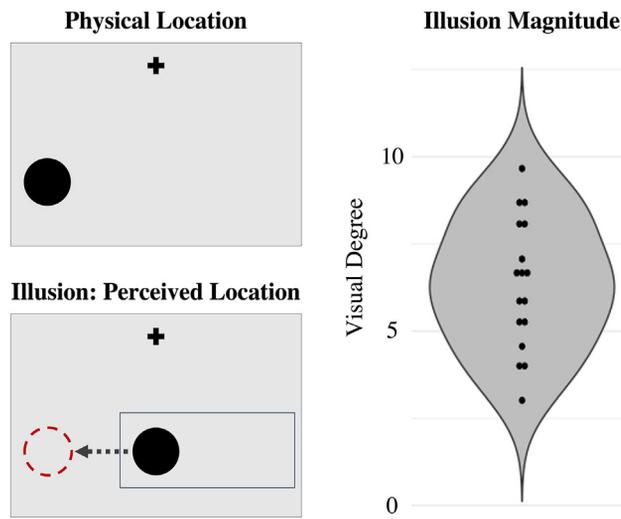https://doi.org/10.1016/j.isci.2023.106800

Figure 1. Illustration of the task displays (left) and reported magnitude of the illusory position shift (right)
When a horizontally moving frame is shown (bottom left), a flashing dot that is presented centrally appears shifted toward the left or right sides. Using the reported illusion magnitude, we constructed the control condition where the dots are physically shifted the same amount (top left). Each participant's pre-EEG session illusion magnitude is plotted on the right.

showing that activation patterns in early visual cortex carry information about the perceived locations.[9] Together, these data indicate that early visual processing is involved in coding the perceived position of objects. However, the recently discovered FIPS illusion is thought to be distinct from these motion-induced illusions, in particular because the magnitude of the perceived position shifts is not affected by the speed of the frame – which is different in other motion-induced illusions, such as the flash-grab effect. Instead, the critical factor that determines FIPS illusion appears to be the distance that the frame travels from one side to the other.[5] Thus, it remains an open question of what the temporal dynamics are that underlie this frame-induced illusory position shift.

In the present study, to tease apart physical and perceived location processing during FIPS, we measured neural activity elicited by small black disks ('probes') that were presented at the exact same physical location (along the central vertical midline) but *perceived* at different, peripheral locations because of a horizontally moving frame, and compared this to probe-elicited neural activity of stimuli that were physically at separate peripheral locations but matched the perceived locations of the illusory shifts. Our results reveal that visually evoked potentials (VEPs) elicited by the probes were sensitive to the physical locations of stimuli starting at ~70 ms, as expected from previous research (e.g.,[10]), whereas later VEPs starting at ~140 ms showed sensitivity to both illusorily perceived and physical location information. Using single-trial multivariate analyses, we further honed in on these effects to test whether similar neural activity patterns underlie these univariate effects. We found reliable within-condition decoding during both the earlier and later time intervals, indicating that our multivariate pattern analyses were sensitive to other aspects of position coding that were not captured in the occipital VEP components. Critically, we also found that the probe locations could be reliably decoded across the physical and perceived locations during the later time interval (140–180 ms) only, but not during the earlier time interval (70–110 ms). Together, this suggests that differential neural processes underlie position coding of illusory and physical locations earlier on in time, but that later neural processes are shared between physical and illusory locations.

RESULTS

Behavior: Assessing individual illusion magnitudes before and after the EEG session

First, we assessed the magnitude of the perceived position shifts in all participants. Participants viewed flashing black disks that were surrounded by a frame that continuously moved back and forth from one side to the other (Figure 1, left bottom). After watching this stimulus, participants were asked to adjust the position of two disks so that they would match the distance between the previously seen disks. This method of adjustment allowed us to estimate the magnitude of the perceived position shift in each
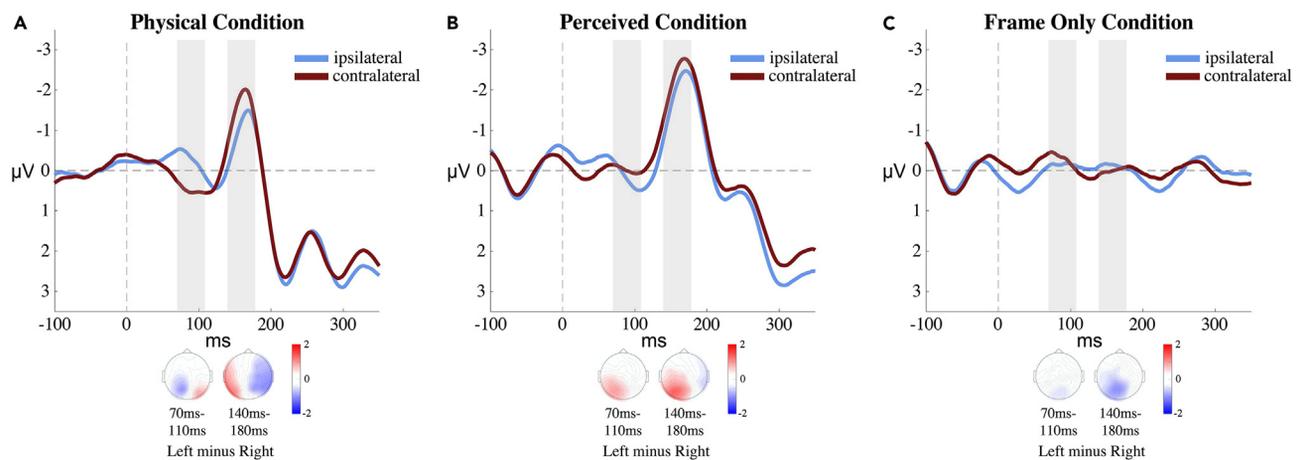
Figure 2. ERP waveforms over parietal-occipital electrode sites with the two time windows of interest highlighted in gray

(A) In the physical condition, both P1 (70-110 ms) and N1 (140-180 ms) components show lateralized effects where amplitudes are larger over the hemisphere contralateral (red) relative to ipsilateral (blue) with respect to the disk location.

(B) In the perceived condition, only the N1 component shows a reliable lateralized effect, with the contralateral waveform showing a larger amplitude than ipsilateral.

(C) These effects cannot be explained by the presence of the moving frame as no lateralized activity was observed there. All waveforms are averaged across eight occipital-parietal electrodes: O1/O2/PO3/PO4/PO7/PO8/P7/P8.

individual participant (for more details, see STAR Methods). Illusion magnitudes were calculated by averaging the distance between the two response disks across three blocks of the behavioral session. Participants completed this adjustment task before and after the EEG session to ensure that participants would still perceive the illusion after being exposed to it for a longer time period. The average illusion magnitude for the pre-EEG sessions was 6.36° of visual angle. Of interest, the average post-EEG session illusion magnitude was significantly reduced to 4.92° ($t(16) = 3.84$, $p = 0.001$; Cohen's $d_z = 0.61$). Most importantly for the current study, both pre- and post-EEG session illusion magnitudes were reliably above 0° (pre-EEG: $t(16) = 13.89$, $p<0.001$; post-EEG: $t(16) = 8.12$, $p<0.001$) demonstrating that participants persistently saw the illusion before and after the EEG session despite the reduction in the magnitudes.

ERPs: Visually-evoked potentials for physical and perceived stimulus locations

There were three conditions in the EEG session: Perceived Location condition (disks are physically at the central location but perceived to appear on the left or right side because of a moving frame), Physical Location condition (disks are physically offset from the center matching the positions of each individual observer's illusion magnitude), and Frame Only condition (only a moving frame is presented without the disks). The display in the Perceived Location condition was identical to the behavioral session where disks were presented every time the moving frame hit either left or right edge of the path. In the EEG session, however, most of the disks were gray placeholders and only 2 or 3 of them per trial were black dots. The black disks were the critical stimuli for the EEG analysis, used to elicit the VEPs. In the Physical Location condition only the black disks were presented, and the Frame Only condition did not have any disks presented. Participants' task was to count the number of black disks (0, 2 or 3) on each trial while fixating their gaze at the top of the screen. For more details, see STAR Methods.

Previous research has demonstrated that brain activity assessed via EEG differs across hemispheres for peripherally presented visual stimuli. Specifically, the P1 and N1 components of the event-related potential (ERP) have been found to be particularly sensitive to stimulus location, often occurring earlier in time and with larger amplitude when recorded over the parietal-occipital cortex contralateral to the stimulus location relative to ipsilateral (e.g.,[10],[11]). Thus, the main question of interest was whether we find similar lateralized effects for both the physical and perceived locations of the probe stimulus. Thus, we first examined the visually evoked potential over occipital scalp sites that was elicited by the disks that were physically located on the left and right sides of the screen. Indeed, as can be seen in Figure 2A, both the P1 (70ms–110 ms) and N1 (140ms–180 ms) components elicited larger amplitudes over the contralateral relative to ipsilateral occipital cortex with respect to the probe's location.
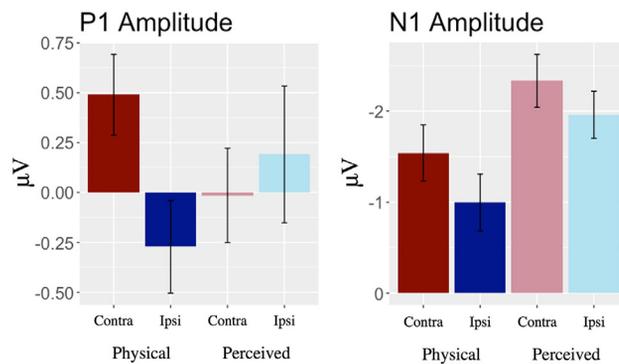
Figure 3. Mean amplitudes plotted for the P1 (70-110 ms; left) and N1 (140-180 ms; right) components separately for the physical and perceived conditions

Error bars represent within-subject standard errors of the mean. For the physical condition (darker colors), both P1 and N1 show a reliable amplitude modulation such that activity recorded over contralateral occipital cortex is larger relative to activity recorded ipsilaterally. For the perceived location condition (lighter colors), only the N1 component shows this lateralized difference in amplitude.

Next, we examined whether the frame-induced perceived left and right locations generate a similar pattern in the ERP. Thus, we compared the waveforms over the hemisphere ipsilateral and contralateral with respect to the *perceived* probe location, although the physical locations were identical and non-lateralized along the vertical meridian (see Figure 1 left bottom). Thus, if early P1 and N1 components are only sensitive to physical, but not perceived location information, we would expect to see no difference between these conditions. As can be seen in Figure 2B, we observed no reliable amplitude difference between the waveforms recorded over the contralateral and ipsilateral hemisphere in the early P1 time period. Of interest, however, we observed larger N1 amplitude over contralateral relative to ipsilateral occipital cortex, resembling the amplitude difference we observed in the physical stimulus condition (see Figure 2).

To test whether these effects were reliable, we conducted 2 × 2 within-subjects ANOVA with disk condition (physical vs. perceived) and laterality (contralateral versus ipsilateral) as factors separately for each of these ERP components. For the P1 component, this comparison yielded a significant interaction ($F_{(1, 17)}$ = 35.476, $p < 0.001$), and no main effects (disk condition: $F_{(1, 17)}$ = 0.007, $p = 0.934$; laterality: $F_{(1, 17)}$ = 4.048, $p = 0.06$). Pairwise comparisons using the t-tests with Bonferroni correction showed that the P1 amplitudes were significantly different in the physical condition but not in the perceived condition (physical condition P1: $t_{(17)}$ = 4.88, $p < 0.001$; perceived condition P1: $t_{(17)}$ = 1, $p = 1.24$). For the N1 component, a 2X2 within-subjects ANOVA yielded a significant main effects of disk condition ($F_{(1, 17)}$ = 6.143, $p = 0.024$) and laterality ($F_{(1, 17)}$ = 10.719, $p = 0.004$), but no significant interaction ($F_{(1, 17)}$ = 0.842, $p = 0.372$, see Figure 3). Given the non-significant interaction, we ran a post-hoc Bayesian ANOVA on N1 amplitudes with the same factors. This analysis yielded substantial evidence for the main effect with both factors combined (BF_{10} = 91.76) and moderate evidence for no interaction (BF_{01} = 2.94), consistent with the previous ANOVA results. Thus, these statistical results indicate that the earlier P1 component was sensitive to the physical location of the disks, whereas the later N1 component was sensitive to both the physical and perceived locations. Of interest, we also observed a main effect of amplitude between the conditions in the later time interval that we did not necessarily anticipate. At this point, we do not know what processes may underlie this main effect. It could be a reflection of overall stronger visual processing during the illusion condition relative to the physical condition, possibly because of stronger feedback from other areas.

To perceive the illusion, a continuously moving frame needs to be presented throughout the perceived location condition that was absent from the physical condition. To ensure that the differences in ERPs were not caused by the moving frame itself, we also analyzed the ERP amplitudes of the frame-only condition where there were no disks presented during the trial. Because there were no disks to reference the laterality when computing the waveforms, we matched this analysis to the analysis used in the perceived condition where the frame being on the left side would imply a disk on the right side of the screen and vice versa. As expected, we found no clear visually-evoked potentials in the same P1 and N1 time windows and no differences between the contralateral and ipsilateral amplitudes in either time
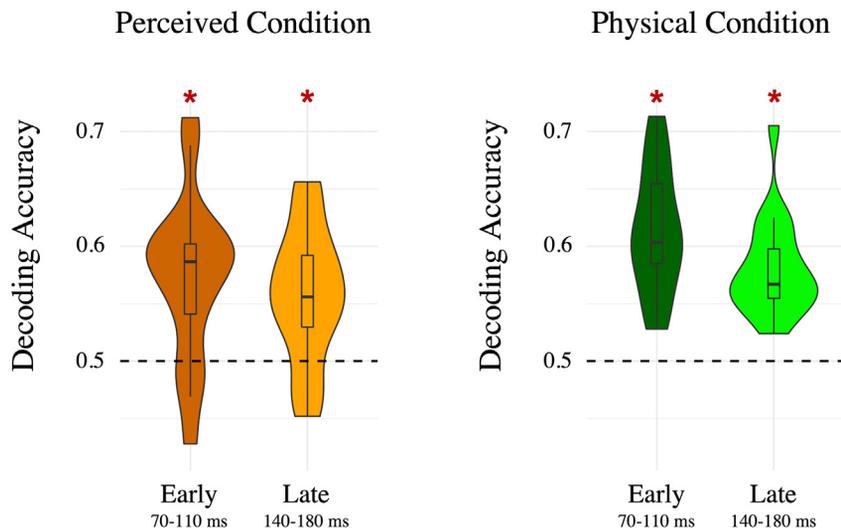
Figure 4. Within-condition decoding results for the perceived condition (left) and physical condition (right)
All conditions and time points resulted in above chance (50%) decoding performance. For the early time interval only, we observed higher decoding accuracy for the physical condition compared to the perceived condition.

window (P1: t(17) = 0.89, p = 0.39; N1: t(17) = 0.77, p = 0.45; see Figure 2C). Thus, the moving frame alone did not produce significant lateralized differences in the ERP.

Neural activity patterns: Multivariate decoding analysis across physical and perceived stimuli

The ERPs provide clear evidence that perceived location information arises later in time than physical location information during FIPS. However, comparing the mean amplitudes of the ipsilateral vs. contralateral waveforms is a relatively coarse measure that tells us little about whether the neural processes underlying the coding of perceived and physical locations are different or shared. Thus, we extended our analyses to examine the neural activity patterns between processing the illusory positions versus the physical positions. To do so, we used a support vector machine (SVM) decoding approach using single-trial EEG activity in each condition. We focused this analysis on the two *a priori* defined time windows of interest that we also used in our ERP analyses. We reasoned that this would increase power and reduce the possibility of false positives, given that single trial data is relatively noisy. Furthermore, we were particularly interested to see whether the differences we observed in the univariate analyses would translate to multivariate decoding as this allows us to test whether the neural activity patterns for perceived and physical locations are shared or not. Thus, our decoding analysis focused on the *a priori* defined early (P1 component, 70ms–110 ms) and late (N1 component, 140ms–180 ms) time windows. We first established a baseline for how well we could decode the left versus right locations of the disks in the physical condition where disks were presented on the left and right side of the screen. For each participant, we trained an SVM to distinguish between the left versus right disk locations based on all single-trial EEG data using the activity pattern over 31 electrodes averaged over each of these two time windows, and then tested this classifier on a leave-one-out hold-out set (see STAR Methods). We then averaged individual's decoding accuracy across subjects and found an average decoding accuracy of 61.3% in the early time window and 58.1% in the later time window, both were well above chance (50%; early time window: t(17) = 8.99, p<0.001; late time window: t(17) = 8.06, p<0.001; see Figure 4).

We then tested how well location information could be decoded in the perceived condition using the same approach. Of interest, we found that decoding accuracies of disk location conditions (left versus right) were significantly above chance level in both the early time window (mean performance 57.3%; t(17) = 4.37, p<0.001) and late time window (mean performance 58.1%; t(17) = 4.18, p<0.001; Figure 4), indicating that the multivariate analysis was sensitive to different aspects of perceived position coding that the univariate ERP analysis did not capture. We also compared these decoding accuracies across conditions, and found that although the early time window resulted in above-chance level decoding in both physical and perceived conditions, the physical location condition showed significantly higher decoding
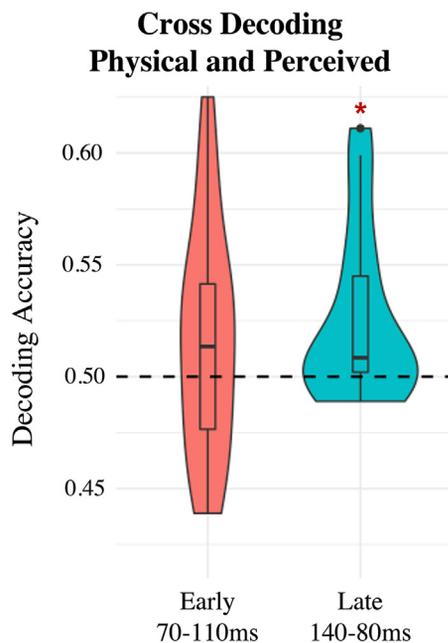
Cross Decoding
Physical and Perceived

Figure 5. Across-condition decoding results
We trained a classifier on the physical probe locations and tested whether it could reliably predict the perceived probe locations. Consistent with the late-arising position information for illusory locations observed in the ERP analysis, we find that we can only reliably cross-decode during the late time window.

performance than the perceived location condition in this early time window (t(17) = 2.61, p = 0.018). This was not the case for the late time window, in which case there was no significant difference in decoding performance between physical and perceived locations (t(17) = 1.65, p = 0.12).

Using this decoding technique we can now ask: Are the illusory locations eliciting the same pattern of activity as the physical locations? For our main across-condition decoding analysis we trained the model on data from the physical condition and tested on a random trial from the perceived condition, reasoning that the physical location condition should give us the most robust and consistent data on single trials, thus serving as a strong training set for our model. In the early time window this cross-decoding performance was not significantly above chance level (mean performance 51.6%; t(17) = 1.39, p = 0.18). However, in the late time window cross-decoding accuracy was significantly above chance level (mean performance 52.5%, t(17) = 2.86, p = 0.01), suggesting that the neural processes that give rise to physical and perceived locations are – at least in part – shared during the late time period (see Figure 5).

For completeness, we also ran the across-condition decoding analysis by training the model on data from the perceived condition and testing on a random trial from the physical condition. The average decoding accuracy for the early time window was 51.4% and not significantly above chance level (t(17) = 0.78, p = 0.44), whereas the average accuracy for the late time window was higher, 52.2%, and marginally non-significant (t(17) = 2.08, p = 0.053). Overall, this replicates the pattern we saw in our primary decoding analysis (late time window resulted in better decoding accuracy compared to early time window), though we did see a decrease in the reliability in which we could cross-decode during the later time window. This may be because of the EEG activity from the perceived condition being overall noisier and more variable relative to the data from the physical condition for a few reasons: First, the illusion condition also contained the presentation of a moving frame, possibly diluting the disk position signal; second, it seems plausible that the perceived illusory positions were not as consistent on each trial but varied slightly from trial to trial and across the experiment relative to the physical positions. Indeed, our behavioral pre- and post-test showed a decrease in the magnitude of the illusion across the experiment. Together, this would mean that our model is trained on a noisier and more variable set of data, which would decrease cross-decoding accuracies. Importantly, the results of this additional decoding analysis were not significantly different from the primary cross-decoding results, however, where we trained the model on data from the physical condition and tested on data from the perceived condition (difference between the two cross-decoding results for the early time window: t(17) = 0.357, p = 0.73, BF_{01} = 3.09; and late time window: t(17) = 0.614, p = 0.55, BF_{01} = 3.01). Thus, overall, the multivariate results indicate that there is a shared pattern of neural activity between processing of illusory and physical locations in the later, but not earlier, time window.

CellPress
OPEN ACCESS

DISCUSSION

Despite much research pointing to the importance of relational position coding to achieve visual stability, how the visual system accomplishes this – and at what stage of processing – remains not fully understood. Using a recently discovered visual illusion where a moving frame induces a paradoxical stabilization of disk positions,[5] we here show that relative position coding arises relatively late in the visual processing stream, but once reached, shares a similar neural code as position coding of objects that were not stabilized because of a frame.

We first examined visually evoked potentials that were elicited by physical peripheral probes (with no frame present) or probes that matched these locations *perceptually* because of a moving frame, but were actually presented at the same physical position along the vertical midline. Our ERP findings showed a clear pattern: the earlier P1 component showed a lateralized effect only for the disks that were physically offset in location, with no reliable difference for perceptually shifted locations (frame-induced position shifts), whereas the later N1 component was sensitive to both physical and illusory disk locations. The P1 and N1 components are well established and long studied visually evoked potentials known to reflect the early visual processes in occipital visual areas (e.g.,[12–16]). The P1 component is typically interpreted as reflecting the initial feed-forward sweep of visual processing, whereas the N1 component is thought to reflect both feed-forward and feedback processing from higher visual and parietal areas. Consistent with this, prior studies have shown that dipoles that are underlying the P1 component are localized to various areas of the striate and extrastriate cortex (i.e., V3, V3a, middle occipital gyrus, and fusiform gyrus), while N1 generators are much more complex, distributed over more diverse areas including extrastriate cortex and also centro-parietal areas (e.g.,[17–24]). Thus, our results suggest that physical location information is coded rapidly in the initial feed-forward sweep of visual information processing (as indexed by the P1), but that later perceptual coding of information – here, the frame-induced stabilization of object locations, arises later in time (as reflected by the N1 component). This timing difference between physical versus illusory information may arise from more recurrent feedback and computations required to achieve visual stability via relational coding.

One potential concern in our design is that P1 and N1 amplitudes could be influenced by the physical difference between the displays of the two conditions: To induce the illusion, a continuously moving frame had to be presented which was not present in the physical condition. Although neither the motion nor the location of the frame would necessarily predict larger P1 or N1 amplitudes at the contralateral side of the perceived disk location (e.g., the frame was always present at the opposite side of the perceived disk location), we also included the moving frame-only condition where no disks were presented on the screen to ensure that the differences in ERP amplitudes were indeed not driven by the presence of the frame. This control condition did not result in any lateralized activity at the two time windows of interests; thus, it is unlikely that the frame itself affected the amplitude differences of the probe-evoked P1/N1 components in the illusory condition. To look into this further, we also examined the overall amplitudes of the P1 and N1 components across the two conditions and found no difference for the P1, but an unexpected reliable difference in amplitude for the N1 component, which was overall higher for the perceived condition compared to the physical condition. This may suggest that the processing of the illusory locations may engage additional neural processes, for example, stronger feedback from higher brain areas that are not present in the physical condition. However, this finding was not part of our *a priori* hypothesis and should be replicated before making strong conclusions. Regardless of the overall amplitude differences, we found no significant interaction of the lateralized effect across the two conditions, indicating that the main comparison of the contralateral versus ipsilateral N1 amplitudes was not affected by the overall difference across the disk conditions.

Using the multivariate decoding analysis, we further investigated whether the neural pattern of activity is sensitive to the physical and illusory locations across different time windows. Results showed that both physical and illusory locations could successfully be decoded during the early time window (70–110 ms) and the late time window (140-180 ms). Although the decoding accuracies are rather low, this is expected given noisy single-trial EEG data, in line with previous studies that have used single-trial EEG activity to decode various task performances (e.g., [25–28]). Despite overall low decoding accuracies, we consistently found decoding accuracies that were reliably above chance level in within-condition decoding. Of interest, we found that decoding performance was significantly above chance level during the early time window for the illusory condition although there were no significantly differentiated waveforms found in the ERP

analysis. This difference illustrates that the multivariate pattern analysis was sensitive to different aspects of neural processing than the targeted ERP analysis that focused on averaged activity at predefined occipital electrode sites, more broadly demonstrating the usefulness of decoding approaches in neural data. Another possibility could be that the constant presentation of the moving frame during the illusory condition was – at least in part – contributing to the decoding results. To directly compare how the neural patterns elicited by the disks compare between processing of the physical and illusory locations, we examined how well the model generalizes from the physical condition to the perceived condition (and vice versa). Results showed that cross-decoding between the two conditions were above chance level during the late time window, but not during the early time window. This suggests that the illusory positions share a neural code with the physical positions later on in time, and that early on in visual processing, these positions are dissociable by the patterns of electrical brain activity.

Our results relate to other studies that investigated the electrophysiological basis of illusory position shifts. In particular, Hogendoorn et al.[8] examined electrophysiological activity during the motion-induced position shift, a form of position shift that is affected by motion of the stimuli, and found that cross-decoding between physical and illusory positions can be successful as early as 81 ms after the stimuli onset, a time point that's analogous to the early time window in our analysis. The different results from Hogendoorn et al.[8] and the current results where we find above chance level cross-decoding only during the late time window can be explained by the differences in the mechanisms underlying motion-induced position shift and FIPS. Behaviorally it has been shown that unlike the motion-induced position shift, FIPS is not affected by low-level stimulus factors such as the speed of the motion or the location of where the motion energy is. It could be that the motion-induced position shift relies more on the low-level visual processes whereas FIPS requires higher level processes that result in shifting the relevant electrophysiological response to a later time point that reflects these illusory positions. It is also true that in the current paradigm the two conditions differ in their stimulus properties (i.e., one has motion energy, one does not), and this might be impacting the decoding performance across the two conditions. However, this seems unlikely to be the main driver of the observed effects as such stimulus-level differences should impact both the early and late time windows, and we selectively find above chance cross-decoding performance for the late time window. Future investigations could aim to control for the differences between the two conditions.

Most broadly, our results suggest that physical location information is coded in the initial feed-forward sweep of visual information processing and that later recurrent processes are involved in the relative coding of object locations to support visual stabilization.

LIMITATIONS OF THE STUDY

The illusion magnitudes we found in the current study are on average smaller in comparison to the previous reports that showed 100% stabilization (illusion magnitude matching the frame's path length; 5). Smaller illusion magnitude in the current study may be explained by the presence of a fixation cross that is necessary in the EEG task to avoid eye movements. Although previous investigations on FIPS claim that having a fixation point does not eliminate the illusion,[5] having a concrete reference point that is also presented at the vertical meridian may produce reduction in the illusion magnitude. Importantly, however, participants reported to consistently perceive the central disks shifted in location before and after the EEG session in the current study.

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
  - Experimental setup and apparatus
  - Experimental procedure
  - Electrophysiological recordings

- QUANTIFICATION AND STATISTICAL ANALYSIS
  - ○ Behavioral data analysis
  - ○ Electrophysiological data analysis

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

REFERENCES

1. Post, R.B., Welch, R.B., and Whitney, D. (2008). Egocentric and allocentric localization during induced motion. Exp. Brain Res. *191*, 495–504.

2. Wertheim, A.H. (1981). On the relativity of perceived motion. Acta Psychol. *48*, 97–110.

3. Morgan, M., Grant, S., Melmoth, D., and Solomon, J.A. (2015). Tilted frames of reference have similar effects on the perception of gravitational vertical and the planning of vertical saccadic eye movements. Exp. Brain Res. *233*, 2115–2125.

4. Asch, S.E., and Witkin, H.A. (1948). Studies in space orientation; perception of the upright with displaced visual fields. J. Exp. Psychol. *38*, 325–337.

5. Özkan, M., Anstis, S., 't Hart, B.M., Wexler, M., and Cavanagh, P. (2021). Paradoxical stabilization of relative position in moving frames. Proc. Natl. Acad. Sci. USA *118*, 1–8.

6. Cavanagh, P., Wexler, M., and Anstis, S. (2020). Frame-induced position shifts. J. Vis. *20*, 607.

7. Cavanagh, P., Anstis, S., Lisi, M., Wexler, M., 't Hart, M., 't Hart, B.M., Shams-Ahmar, M., and Saleki, S. (2022). Exploring the frame effect. J. Vis. *22* (5), 5–13.

8. Hogendoorn, H., Verstraten, F.A.J., and Cavanagh, P. (2015). Strikingly rapid neural basis of motion-induced position shifts revealed by high temporal-resolution EEG pattern classification. Vision Res. *113*, 1–10. https://doi.org/10.1016/j.visres.2015.05.005.

9. Kohler, P.J., Cavanagh, P., and Tse, P.U. (2017). Motion-induced position shifts activate early visual cortex. Front. Neurosci. *11*, 168.

10. Woldorff, M.G., Fox, P.T., Matzke, M., Lancaster, J.L., Veeraswamy, S., Zamarripa, F., Seabolt, M., Glass, T., Gao, J.H., Martin, C.C., and Jerabek, P. (1997). Retinotopic organization of early visual spatial attention effects as revealed by PET and ERPs. Hum. Brain Mapp. *5*, 280–286.

11. Luck, S.J., and Hillyard, S.A. (1994). Spatial filtering during visual search: evidence from human electrophysiology. J. Exp. Psychol. Hum. Percept. Perform. *20*, 1000–1014.

12. Heinze, H.J., Mangun, G.R., and Hillyard, S.A. (1990). Visual event-related potentials index perceptual accuracy during spatial attention to bilateral stimuli. Psychophysiol. Brain Res. 196–202.

13. Hillyard, S.A., and Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. Proc. Natl. Acad. Sci. USA *95*, 781–787.

14. Luck, S.J., Hillyard, S.A., Mouloua, M., Woldorff, M.G., Clark, V.P., and Hawkins, H.L. (1994). Effects of spatial cuing on luminance detectability: psychophysical and electrophysiological evidence for early selection. J. Exp. Psychol. Hum. Percept. Perform. *20*, 887–904.

15. Mangun, G.R., and Hillyard, S.A. (1995). Mechanisms and models of selective attention. Electrophysiology of Mind: Event-Related Potentials and Cognition, pp. 40–85.

16. McDonald, J.J., Störmer, V.S., Martinez, A., Feng, W., and Hillyard, S.A. (2013). Salient sounds activate human visual cortex automatically. J. Neurosci. *33*, 9194–9201.

17. Eimer, M. (1998). Mechanism of visuospatial attention: evidence from event-related brain potentials. Vis. Cognit. *5*, 257–286.

18. Hashimoto, T., Kashii, S., Kikuchi, M., Honda, Y., Nagamine, T., and Shibasaki, H. (1999). Temporal profile of visual evoked responses to pattern-reversal stimulation analyzed with whole-head magnetometer. Exp. Brain Res. *125*, 375–382.

19. Hoshiyama, M., and Kakigi, R. (2001). Effects of attention on pattern-reversal visual evoked potentials: foveal field stimulation versus peripheral field stimulation. Brain Topogr. *13*, 293–298. https://doi.org/10.1023/A:1011132830123.

20. Nakamura, M., Kakigi, R., Okusa, T., Hoshiyama, M., and Watanabe, K. (2000). Effects of check size on pattern reversal visual evoked magnetic field and potential. Brain Res. *872*, 77–86. https://doi.org/10.1016/S0006-8993(00)02455-0.

21. Brecelj, J., Kakigi, R., Koyama, S., and Hoshiyama, M. (1998). Visual evoked magnetic responses to central and peripheral stimulation: simultaneous VEP recordings. Brain Topogr. *10*, 227–237.

22. Seki, K., Nakasato, N., Fujita, S., Hatanaka, K., Kawamura, T., Kanno, A., and Yoshimoto, T. (1996). Neuromagnetic evidence that the P100 component of the pattern visual evoked response originates in the bottom of the calcarine fissure. Electroencephalogr. Clin. Neurophysiol. *100*, 436–442. https://doi.org/10.1016/s0921-884x(96)95098-5.

23. Di Russo, F., Martínez, A., Sereno, M.I., Pitzalis, S., and Hillyard, S.A. (2002). Cortical sources of the early components of the visual evoked potential. Hum. Brain Mapp. *15*, 95–111.

24. Sun, R., Sohrabpour, A., Worrell, G.A., and He, B. (2022). Deep neural networks constrained by neural mass models improve electrophysiological source imaging of spatiotemporal brain dynamics. Proc. Natl. Acad. Sci. USA *119*. e2201128119.

25. Brady, T.F., Alvarez, G.A., and Störmer, V.S. (2019). The role of meaning in visual memory: face-selective brain activity predicts memory for ambiguous face stimuli. J. Neurosci. *39*, 1100–1108.

26. Leydecker, A., Biebmann, F., and Fazli, S. (2014). Single-trials ERPs predict correct answers to intelligence test questions. International Workshop Pattern Recognition in Neuroimaging (IEEE Computer Society), pp. 1–4.

27. Noh, E., Herzmann, G., Curran, T., and de Sa, V.R. (2014). Using single-trial EEG to predict and analyze subsequent memory. Neuroimage *84*, 712–723.

28. Höhne, M., Jahanbekam, A., Bauckhage, C., Axmacher, N., and Fell, J. (2016). Prediction of successful memory encoding based

on single-trial rhinal and hippocampal phase information. Neuroimage *139*, 127–135.

29. Brainard, D.H. (1997). The psychophysics toolbox. Spat. Vis. *10*, 433–436.

30. Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. *10*, 437–442.

31. Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods *134*, 9–21.

32. Lopez-Calderon, J., and Luck, S.J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials.

Front. Hum. Neurosci *8*, 213. https://doi.org/10.3389/fnhum.2014.00213.

33. Cox, D.D., and Savoy, R.L. (2003). Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage *19*, 261–270.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESROUCE	SOURCE	IDENTIFIER
Deposited data		
EEG and behavioral data of 18 human observers	This paper	Mendeley Data (https://data.mendeley.com/datasets/9fny432trn)
Software and algorithms		
MATLAB R2021b	MathWorks, USA	https://www.mathworks.com/products/matlab.html
EEGLAB toolbox	Delorme & Makeig, 2004	https://doi.org/10.1016/j.jneumeth.2003.10.009
ERPLAB toolbox	Lopez-Calderon & Luck, 2014	https://doi.org/10.3389/fnhum.2014.00213
RStudio 1.3.1093	RStudio, PBC	https://www.rstudio.com/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Yong Hoon Chung (yong.hoon.chung.gr@dartmouth.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data files for the behavioral tasks and EEG recordings are freely available on Mendeley (Mendeley Data: https://data.mendeley.com/datasets/9fny432trn).

- This paper does not report original code.

- 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

Eighteen participants were recruited from Dartmouth College, all reported normal or corrected-to-normal vision and were between 18 and 28 years of age. All participants gave informed consent as approved by the Committee for the Protection of Human Subjects at Dartmouth College and were paid $20/hr or 1 class credit/hr for their time. All participants included in the study reported seeing the illusion on the demo video, however, three participants who did not show more than 2° of visual angle of illusion on the initial behavioral session were excluded from the study. The value of one participant's illusion magnitude was not saved correctly, resulting in 17 total participants for the behavioral analysis.

METHOD DETAILS

Experimental setup and apparatus

Participants were seated approximately 37 cm in front of a 24-in computer monitor (1920 × 1080) in a dark, electrically shielded chamber. Stimuli were presented on a white background via the Psychophysics Toolbox[29,30] in MATLAB. In both the behavioral and EEG sessions, a small black fixation cross (0.4° × 0.4° of visual angle) was presented at 7° above from the center of the screen. On each trial, a rectangular gray frame (18.5° width x 3° height x 0.15° thickness; exact color value varied across participants, average [RGB: 99 99 99]) was presented at 12° below the center of the screen. During the trial, this frame moved left and right at a speed of 2.6° (60 pixels) per frame within the bounded range of 30° left and right from the center. When the frame made contact with the left or right border, a faint gray or black disk (2.4° radius) appeared 12° below the center of the screen. For the EEG session, these disks were presented on the vertical meridian (Perceived Location condition), or their positions were individually adjusted towards

the left/right sides of the screen based on the individual's magnitude of the illusion (Physical Location condition). In the behavioral part of the experiment, participants only completed the illusory trials (Perceived Location condition) and, after each trial, were instructed to adjust the distance between two response disks to match the locations that they had perceived during that trial (see experimental procedure).

Experimental procedure

Participants first completed the behavioral session and then went on to the EEG session. In the behavioral session, the magnitude of FIPS illusion was measured for each participant. Participants were instructed to maintain their gaze at the fixation cross located at the top of the screen throughout the entire duration of each trial. On each trial, a frame was presented at the bottom of the screen and continuously moved horizontally left and right along a preset path at the speed of 2.6° per 8.3 ms. The frame stopped when it reached the left or right bound of the path, and a disk appeared at the bottom of the screen on the vertical meridian for 100 ms. Each trial consisted of 10 disks flashing. In most cases the disks were a faint gray color ([RGB: 190 190 190]) that mostly blended in with the background. At two random intervals, a black disk was presented: one when the frame reached the left edge, and one when the frame reached the right edge. We used these different disk colors to avoid overlap in the ERPs elicited by the disks during the EEG session while maximizing the illusion magnitude (see more details below). After the illusion display, participants reported the locations of the previously seen two black disks by adjusting the locations of two response probes. These response probes were first presented at the vertical meridian (at the bottom of the screen at center width), just as during the illusion itself, and as participants pressed the 'c' key to move them further apart or the 'm' key to move them closer together. When participants felt confident that the response probes were adjusted to reflect the perceived distance between the two black disks, participants pressed the spacebar and moved onto the next trial. Participants completed 12 trials per block and repeated 3 blocks. Illusion magnitudes were averaged across all three blocks.

The EEG session consisted of 3 conditions: Perceived Location condition (disk+frame), Physical Location condition (disk-only), and Frame Only condition. The display in the EEG session was identical as in the behavioral session: In the Perceived Location condition a moving frame and disks were presented every time the frame hit either left or right edge of the preset path. Dots were presented 10 times per trial with most of them being faint gray placeholders and only 2 or 3 black dots at random intervals. The black disks were the critical stimuli for the EEG analysis: ERPs were time-locked to these high-contrast black disks. The faint gray disks were presented with the sole purpose to strengthen the perception of the illusion which is more apparent with a continuous stream of disk presentations, while not eliciting strong visually-evoked potentials due to their low contrast. In the Physical Location condition, only the black disks were presented at the locations of each individual's illusion magnitude. This means that if the participant reported 5° of illusion, then the disks were presented 5° left or right from the center width. This effectively simulates the illusory locations of each of our participants. The faint gray disks were omitted in the disk-only condition in order to eliminate the expectancy of the disk location based on oscillatory disk presentation. In the Frame Only condition, only the moving frame and faint gray placeholder disks were presented. At the intervals where the black disks would have shown up, no disks were presented on the screen. After each trial participants reported how many black disks were presented during the trial (0, 2, or 3) and received feedback on their response. The purpose of the task was to ensure that participants would be attending to the disks at the bottom of the screen. Participants completed the task until the end of the scheduled session, which resulted on average of 602 total trials per participant (ranging from 488 to 710 trials). Average accuracy of the disk-counting task was 98% correct.

After the EEG session, three more blocks of the behavioral session were run to compare how the illusion magnitudes changed over time.

Electrophysiological recordings

EEG was recorded continuously from 32 Ag/AgCl electrodes mounted in an elastic cap and amplified by an ActiCHamp amplifier (BrainProducts, GmbH). Electrodes were arranged according to the 10–20 system. The horizontal electrooculogram (HEOG) was recorded from two additional electrodes placed on the external ocular canthi which were grounded with an electrode placed on the neck of the participant. All scalp electrodes were referenced to the right mastoid online and were digitized at 500 Hz. EEG data was filtered with a bandpass of 0.01–112.5 Hz online.

QUANTIFICATION AND STATISTICAL ANALYSIS

Behavioral data analysis

Each participant's illusion magnitude measures were averaged across 3 blocks for pre-EEG session and post-EEG session separately. Illusion magnitudes were then compared against 0 using single-sample t-tests. Pre-EEG and post-EEG magnitudes were compared using a two-sample t-test. Statistical analyses were performed in R Studio.

Electrophysiological data analysis

Data processing was performed using EEGLAB[31] and ERPLAB[32] toolboxes in MATLAB (The MathWorks, Natick, MA) and custom-written scripts. Data were epoched from −100 ms to 450 ms with respect to the onset of the black disk (or the equivalent time period for the frame only condition). A semiautomatic procedure was performed to remove any EEG epochs that were contaminated by artifacts such as eye movements and blinks. For each participant, each epoch was visually inspected to check the automated procedure. Artifact-free data were re-referenced to the average of the left and right mastoids.

ERP analysis: For the ERP analysis, waveforms were averaged separately with respect to the left and right disk locations and were then collapsed across disk location and hemisphere of recording to obtain waveforms recorded ipsilaterally and contralaterally relative to the physical or perceived locations. For the frame-only condition, the same procedure was performed as for the perceived location condition where the dot location was assumed to be left or right with respect to the frame location. ERPs were low-pass filtered (half-amplitude cutoff at 25Hz; slope of 12dB/octave) to remove high-frequency noise. Mean amplitudes for each participant and condition were measured with respect to a 100-ms prestimulus period (−100 to 0 ms from disk onset, or equivalent time period for frame only condition). The exact time windows and electrode sites for each ERP analysis were chosen *a priori* based on previous research and matched across all analyses. Our main analysis was focused on the neural activity during the P1 (70ms–110 ms) and N1 (140ms–180 ms) time windows. Both ERP components were measured at the same eight parietal-occipital electrode sites (O1/O2/PO3/PO4/PO7/PO8/P7/P8). Amplitudes of ipsilateral and contralateral ERPs were then compared using 2 × 2 ANOVA with conditions and laterality as factors, followed up by pairwise t-tests for significant interactions.

Multivariate decoding analysis: For the decoding analysis, we used single-trial EEG activity averaged either over the early time window of the P1 (70-110 ms) or the later time window of the N1 component (140-180 ms). For each participant, we trained a classifier to distinguish between the left and right locations of the stimulus based on a support vector machine (SVM) decoding approach.[33] The input feature vector in each case consists of 31 features: averaged activity of the 40-ms time window from each of 31 electrodes. We then applied a leave-one-out classification procedure where we trained the model using all available data within-participant (all trials we had for a given subject in the condition we trained on) except one trial from each location (left or right), and tested on the randomly chosen left-out trial data. This procedure was iterated for 500 times for each participant and trials were reshuffled randomly each time, and decoding accuracy was calculated by averaging across these 500 iterations for each participant. We applied this method separately for each time window and condition: Perceived Location (disk+frame), Physical Location (disk-only), and Frame Only. Then, to test whether the neural activity pattern is similar across the conditions, we tested whether training in one condition would generalize to another condition. Thus, for each participant, we trained a model on all trials from the disk-only condition (physical location condition) and tested on a random trial of the disk+frame condition (perceived location condition). We also ran across-condition decoding in the opposite direction: training a model on all data from the disk+frame condition (perceived location condition) and testing on the disk-only condition (physical location condition). In each decoding condition, we then performed one-sample t-tests versus chance performance (50%) to assess the statistical significance of the decoding accuracy.