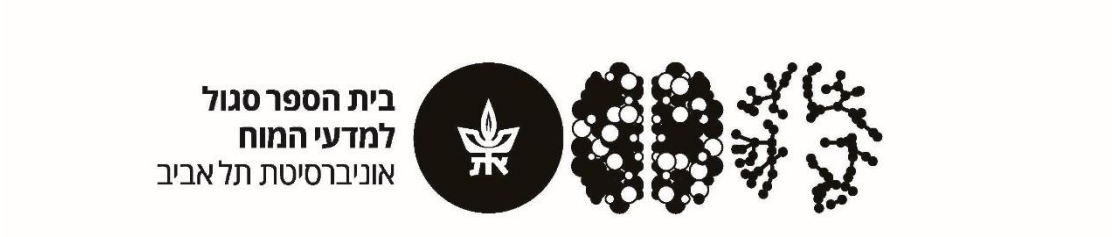
Sagol School of Neuroscience
School of Psychological Sciences

The influence of active tracing on shape representation in visual pathways

By
Guy Baratz

The thesis was carried out under the supervision of
Prof. Roy Mukamel

September 2023



בית הספר סגול למדעי המוח
ביה"ס למדעי הפסיכולוגיה

ההשפעה של שרטוט אקטיבית על ייצוג צורות בנתיבים חזותיים

מאת

גיא ברץ

החיבור בוצע בהנחייתו של

פרופ' רועי מוכמל

ספטמבר 2023

Abstract

How a stimulus is perceived is contingent on the physical properties of the stimulus and the neural state of the individual perceiving it. Specifically, research has demonstrated that voluntary actions modulate neural responses and the corresponding perception of their sensory consequences relative to otherwise physically identical stimuli from an external source. Although the functional significance of such modulations is unknown, one possibility is that they sharpen the neural representation of expected sensory outcomes. We used fMRI to investigate the neural representation of visual shapes during active shape tracing and passive viewing of similar dynamic shape traces. We hypothesized that motor engagement facilitates the distinction between the neural representation of different shapes. We operationalized this by examining the decoding accuracy levels of fMRI activity in visual regions evoked by two different visual shapes. In the scanner, participants completed two experimental conditions: right-hand active tracing of two distinct shapes using an MR-compatible drawing board and viewing dynamic shape traces of those stimuli collected outside the scanner. Additionally, participants completed localizers to identify visual regions sensitive to shapes and to visual motion.

The findings (n=14) demonstrated that the active tracing condition resulted in higher shape classification levels in visual regions relative to observing shape traces. Additionally, in the passive trace-viewing condition the right hemisphere of the occipital pole had higher mean classification accuracies than the left hemisphere, while no such differences were found in the active tracing condition. This pattern of results is compatible with the notion that the neural representation of visual shapes is sharpened by motor engagement, particularly in the visual regions ipsilateral to the active motor cortex.

תקציר

האופן שבו גירוי מסויים נתפס תלוי בתכונות הפיזיקליות של הגירוי ובמצב העצבי המקדים של האדם התופס אותו. באופן ספציפי, מחקרים הראו כי פעולות רצוניות והתוצאה החושית שלהן עוברים מודולציה כתלות במקור הגרוי, באם נוצר באופן עצמאי או על ידי גורם חיצוני. למרות שהמשמעות התפקודית של מודולציות אלה אינה ידועה, אפשרות אחת היא שהן מחדדות את הייצוג העצבי של התוצאה החושית הצפויה.על כן, השערת המחקר הייתה כי מעורבות מוטורית תחדד את ההפרדה בין הייצוגים העצביים של צורות השונות במחקר הנוכחי, בחנו את השערה זו על ידי השוואת רמות הניבוי של הצורות השונות בהתבסס על דפוס הפעילות הנוירונלית באזורים ויזואלים בשני תנאים: אקטיבי ופאסיבי. במהלך סריקת ה fMRI הנבדקים השלימו מטלות התואמות לכל אחד משני התנאים במחקר: תנאי אקטיבי, במהלכו נבדקים שרטטו שתי צורות שונות ביד ימין, באצמעות טאבלט המותאם לשימוש בסורק ה MRI ותנאי פאסיבי, במהלכו נבדקים צפו בהתהוות קווי המתאר של הצורות על המסך (אשר הוקלטו בשלב מקדים לסריקת הfMRI). בנוסף, הנבדקים השלימו שתי סריקות בלתי תלויות לצורך זיהוי האזורים הראייתיים המעורבים בתפיסת גירוי ויזואלי דינמי וצורות.

התוצאות מ 14 נבדקים הראו כי בתנאי האקטיבי יכולת הניבוי של הצורות השונות על בסיס הייצוג העיצבי באזורים ראייתיים הינה גבוהה יותר לעומת התנאי הפאסיבי. יתרה מכך, התוצאות מצביעות על הבדל בין ההמיספרות כך שבתנאי הפאסיבי יכולת הניבוי של הצורות היתה גבוהה יותר בקוטב האוקסיפיטלי הראייתי של המיספרה ימין מאשר שמאל. לא נראה הבדל שכזה בין ההמיספרות בתנאי האקטיבי. תוצאות אלה תואמות עם ההנחה כי חדות הייצוג העצבי של צורות ויזואליות הינה גבוה יותר כאשר ישנה מעורבות של פעולה מוטורית, במיוחד באזורי הראייה האיפסילטרלים לקליפת המוח המוטורית הפעילה.

Contents

Introduction

Perception-Action Coupling

Physical reality is not simply recreated through perception but is mediated by an observer's context. This can be understood by way of the phenomenon of simultaneous contrast, in which the darkness of a gray patch is perceived as darker when surrounded by a light background and lighter when surrounded by a dark background (Helmholtz & Southall, 1925; Hering, 1964). Illusions, such as the Ebbinghaus illusion, in which a circle's size is perceived differently based on the circle's surrounding, are another example of contextual mediation of one's perception of physical reality (Massaro & Anderson, 1971). An individual's ability to act on the outside world also influences how they perceive it. This idea was central to Gibson's theory of affordances, in which he states, "the affordances of the environment are what it offers the animal, what it provides or furnishes" (Gibson, 2014). Essentially, affordances represent an individual's capacity to act on the environment (Gibson, 1977). Consider the experience of being placed at the start of a challenging ski trail with a set of skis. Skis offer novice skiers no help descending the slope, instead, they are burdensome if they must be carried to the end of the trail. In contrast, they offer expert skiers the fastest route to descend.

Perception is also altered by an individual's capacity for action. For example, when throwing a heavy ball at a target, people perceive the target to be farther away than when throwing a light ball (Witt et al., 2004). Similarly, reach-extending tools make objects seem closer to perceivers than when reaching for the object without a tool (Witt et al., 2005). Furthermore, an individual's neural activity affects how the external world is perceived and processed, as demonstrated by research revealing that microstimulations of the brain can alter motion perception (Salzman et al., 1990).

Sensory Modulation

In addition to serving as the mechanism through which individuals can alter their environment, voluntary action influences how individuals perceive the world around them and process incoming sensory information (Hughes & Waszak, 2011). Original

findings regarding the influence of voluntary action on sensory processing indicated that voluntary action led to a suppressed neural response compared to externally generated stimuli, leading to the term sensory attenuation. However, in recent years, there have been reports of amplification under certain circumstances (Reznik & Mukamel, 2019). As such, henceforth the influence of motor engagement on perception will be referred to as sensory modulation. This phenomenon has been demonstrated empirically across multiple senses, including touch (Ackerley et al., 2012; Blakemore et al., 1998), vision (Buaron et al., 2020; Cardoso-Leite et al., 2010; Dewey & Carr, 2013; Hughes & Waszak, 2011; Lubinus et al., 2022; Mifsud et al., 2016; Ody et al., 2023) audition (Mifsud et al., 2016; Reznik et al., 2021; Reznik, Henkin, et al., 2015; Reznik et al., 2018), and proprioception (Gritsenko et al., 2007; Laufer et al., 2001).

There has been evidence of sensory modulation in the visual system on behavioral, physiological, and neural levels. For example, Dewey and Carr (2013) found that when self-generated visual motion was congruent with the direction of voluntary action, it was perceived as slower than when the same visual motion was externally generated. On the physiological level, it has been demonstrated that self-generated stimuli result in increased pupil size (Lubinus et al., 2022). Visual stimuli generated by voluntary actions were shown to modulate event-related potentials measured by EEG compared to those evoked by externally generated stimuli (Hughes & Waszak, 2011; Mifsud et al., 2016; Ody et al., 2023; Schafer & Marcus, 1973). Modulated neural activity due to voluntary action has also been identified using fMRI. In multiple brain structures, including the visual and somatosensory cortex, cerebellum, basal ganglia, and thalamus, visual stimuli generated by voluntary action were associated with earlier and shorter blood oxygenation level dependent (BOLD) responses (Kavroulakis et al., 2022). Furthermore, there is evidence that voluntary action encodes additional levels of information, as demonstrated by research showing behavioral and neural modulations based on the lateral relationship between the hand that generates a stimulus and the stimulated visual field (Buaron et al., 2020).

The leading theories explaining the neural mechanisms underlying the phenomenon of sensory modulation were proposed by Sperry (1950), who described such a mechanism as the corollary discharge, while Von Holst (1954) termed it the efference

copy. These terms are often used interchangeably, but slight differences exist between them. Corollary discharges refer to motor-related timing signals that influence sensorimotor processing, while efference copies are signals that cancel predictable reafferent information (Fukutomi & Carlson, 2020). Neural signals, believed to relate to an organisms capacity to distinguish between the sensory consequences of their own actions from those from external sources, have been identified across the animal kingdom (Crapse & Sommer, 2008). The neuroanatomical source of the efference copy has been attributed to multiple regions in the central nervous system, including the cerebellum (Blakemore et al., 1998), and regions upstream of the primary motor cortex such as the supplementary motor area (Haggard & Whitford, 2004; Voss et al., 2006).

The Forward Model

The capacity to attribute agency to our own actions is critical for effective functioning (Subramaniam et al., 2018). Nevertheless, incoming sensory signals do not contain inherent information that differentiates between stimulus source (self or other). For example, somatosensory information travels through the same sensory pathways regardless of whether it results from voluntary or external stimuli (Miall & Wolpert, 1996). The efference copy is likely to serve as a neural mechanism capable of accomplishing this type of differentiation. The forward model is a prominent theory explaining how the efference copy can serve in motor control (Miall & Wolpert, 1996; Wolpert & Ghahramani, 2000). According to this model, a motor command generates two parallel processes: an action (the motor commands sent to the relevant effector) and an efference copy that predicts the sensory consequences of that action. Using this prediction, it is possible to compare the intentional consequences of motor actions with the actual sensory feedback, known as reafference (see Figure 1). This comparison can then be applied to various aspects of motor execution, including rapid online movement adaptation, sensory modulation, and a sense of agency (Welniarz et al., 2021).

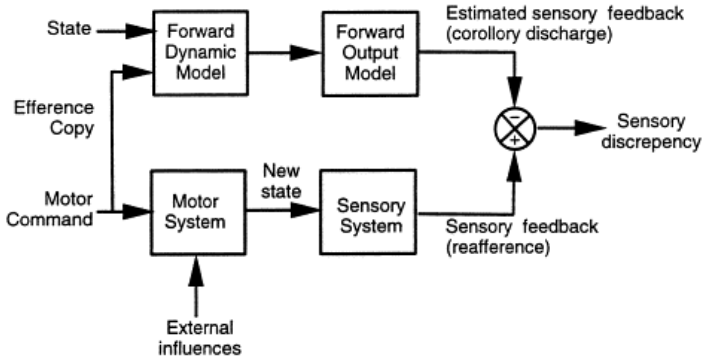


Figure 1: The forward model (Miall & Wolpert, 1996)

Indeed, in pathological conditions such as schizophrenia, it has been shown that on both the behavioral and neural levels, an individual's capacity to distinguish between the sensory consequences of their own actions and those caused by the external world, is compromised (Ford et al., 2001; Shergill et al., 2005; Shergill et al., 2014). The fact that no inherent information enables differentiation between self-generated and externally generated actions, coupled with evidence that the inability to distinguish between the sensory consequences of voluntary action and externally generated stimuli is pathological, lends credence to the notion that humans are capable of differentiating between voluntary action and externally generated stimuli by means of an internal mechanism like the efference copy.

The Forward Model as a Mechanism for Neuronal Sharpening

A common finding in the literature is that voluntary action results in attenuated responses on both perceptual and neural levels (Blakemore et al., 1998). By comparing the predictions of sensory consequences with the actual reafferent signals, a canceling of the redundant prediction occurs, potentially resulting in a discrepancy between the predicted and actual sensory information. One intuitive example of the cancellation effect of voluntary action occurs when one is unable to elicit the same response when tickling oneself as when being tickled by someone else. A classic study by Blakemore et al. (1998) demonstrated the neural basis of this phenomenon by showing that externally generated tickling resulted in more activity in the somatosensory cortex than self-generated tickling. However, recent evidence suggests that the canceling effect of voluntary action is not universal. For example,

Reznik et al. (2014) found that self-generated sounds induced bilateral enhancement of neural signals in the auditory cortex as compared to externally generated sounds. A subsequent study by Reznik, Henkin, et al. (2015) found voluntary action to enhance perceived loudness when the intensity of self-generated sounds approached the hearing threshold. In an EEG study, Mifsud et al. (2016) demonstrated that voluntary action results in suppressed N1 amplitudes in the auditory domain, while the N145 component in the visual domain is strengthened.

An alternative to the cancellation theory suggests that the efference copy sharpens the sensory neural response rather than suppressing it. This theory helps reconcile the discrepant findings regarding the sensory consequences of voluntary action. Neurons in sensory regions of the brain, fire selectively in response to specific features of stimuli (Jia et al., 2010). The neuronal tuning curve, which illustrates how the firing rate of a neuron changes as a result of the stimulus parameters, is the standard tool for analyzing the feature encoded by a neuron (Butts & Goldman, 2006) (see Figure 2). Tuning curves are sharpened when their slopes steepen, and their widths decrease increasing their response selectivity. A sharpened tuning curve is believed to encode information more efficiently (Seriès et al., 2004).

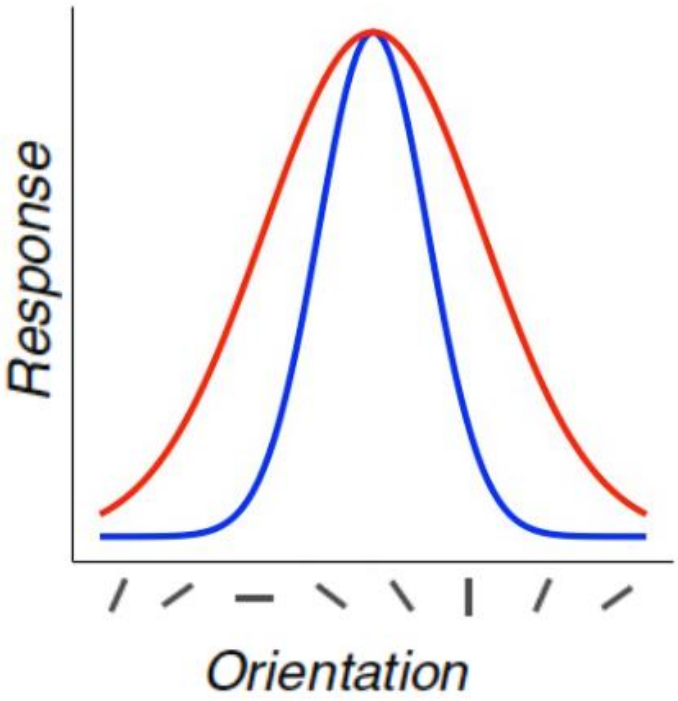


Figure 2: Schematic illustration of neuronal tuning curves to demonstrate orientation selectivity of two neurons in the visual cortex. The response represented by the red Gaussian curve depicts a neurons that is less selective, while the response shown by the blue Gaussian curve represents a neurons more sensitive to deviation from the optimal orientation as indicated by the narrow tuning curve. It is possible that during motor engagement, the tuning curve of individual neurons is sharpened, as demonstrated by the blue Gaussian curve. Adapted from Ling et al. (2015).

Bayesian accounts of perception posit that factoring in prior expectations will provide an advantage while making predictions regarding incoming sensory information (Yuille & Kersten, 2006). From a neural perspective, this can be done through the excitation of expected sensory inputs or the suppression of unexpected ones (Summerfield & De Lange, 2014). Indeed, evidence exists that prior expectation regarding the orientation of a visual stimulus reduces neural activity in the primary visual cortex while also increasing stimulus decoding accuracy when using multivoxel pattern analysis (Kok et al., 2012). This suggests that congruence between expectations and actual sensory input may sharpen neural representation (see Figure 3). The process of neuronal inhibition can also enable more precise filtering of irrelevant neural activity, allowing only the most relevant information to be transmitted to the brain (Isaacson & Scanziani, 2011). This notion helps explain why enhanced perceptual prediction is often paired with suppressed neural activity.

Building on the understanding of how congruence between prior expectation and sensory input influences neural activity, Yon et al. (2018) used fMRI to investigate the sharpening hypothesis by asking participants to abduct either their index or little finger while watching an avatar perform actions either congruent or incongruent to theirs. In this study, decoding accuracies for participant hand actions were higher in early visual regions when visual information was congruent with the action performed, while BOLD responses were decreased in voxels tuned away from the predicted outcome. In addition to corresponding with the aforementioned Bayesian accounts of perception, these findings can help reconcile the inconsistent results concerning the influence of motor engagement on the neural activity in sensory regions. These findings indicate that the nature of the neural response (amplification or suppression) is contingent on the relative influence of a task on the neuronal

populations. The inconsistency observed in past studies may arise from the dynamic interplay between the suppression of one population of neurons and the amplification of another. Essentially, the difference lies in the balance between the amplification and suppression of neuron populations as well as the manner in which measurement devices are able to capture these differences. For instance, relative to EEG, in fMRI, the differences in activity may be more precisely discerned between neuronal populations due to its superior spatial resolution (Glover, 2011).

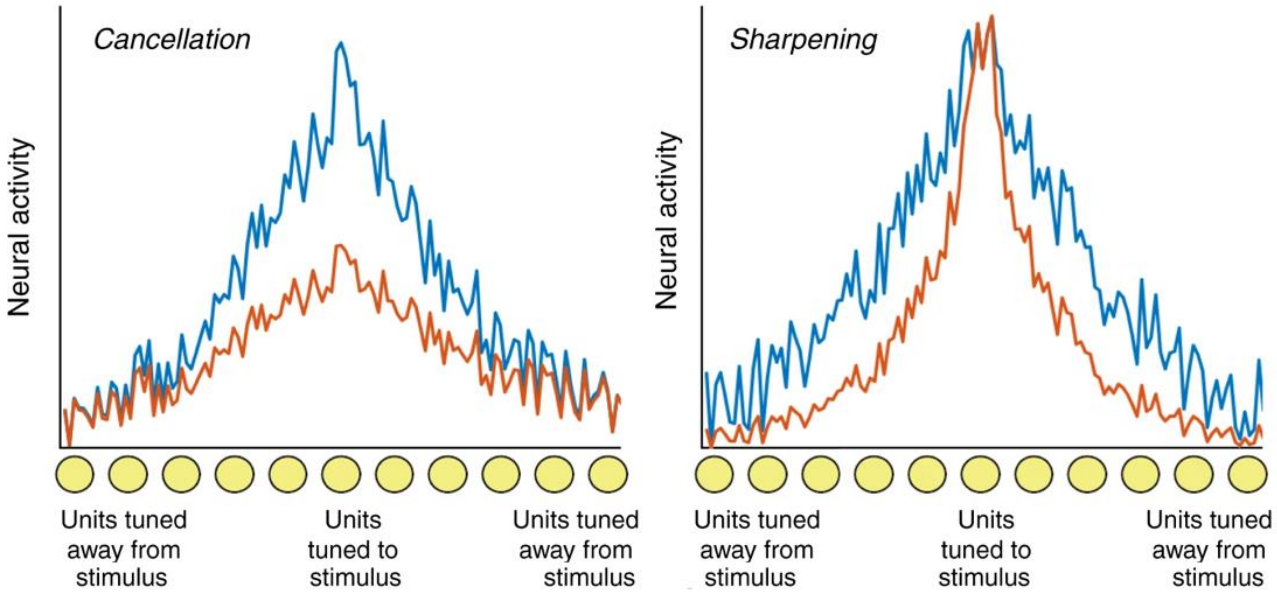


*Figure 3**:** A schematic illustration contrasting a cancellation and sharpening models. Left: The cancellation model suggests that motor engagement suppresses neural activity related to the anticipated action outcome (red) relative to neural responses due to externally generated stimuli (blue). Right: The sharpening model suggests that motor engagement suppresses neural activity unrelated to the expected stimuli but not those related to the expected stimuli (red), relative to neural responses due to externally generated stimuli (blue). Adapted from Yon et al. (2018).*

Research goals and hypotheses

The current study examines whether motor engagement sharpens the neural representation of visual shapes in the visual cortex. We hypothesize that motor engagement will sharpen the neural representation of visual shapes in the visual cortex. To test this hypothesis, we conducted an fMRI study where participants either

actively traced shapes on an MRI-compatible tablet with real-time feedback on their tracing progress or passively observed traces they had previously completed before entering the MRI. We expected to observe higher shape classification accuracies in the active tracing condition within visual regions of the brain, even though the visual information presented in both conditions was similar.

Methods

Participants

We recruited 22 healthy, right-hand dominant participants with normal or corrected to normal vision and normal hearing (15 females, mean age 24.9, range 19-34 years). The study adhered to the guidelines approved by the ethics committee at Tel-Aviv University. All participants provided written informed consent to participate in the study and were provided compensation for their time.

Materials and Procedure

Participants engaged in two experimental sessions – a behavioral session followed by an fMRI scanning session. During the behavioral session, participants were presented with one of two shape templates and instructed to trace them as accurately as possible within eight seconds (slow trials were interrupted and discarded). Tracing was performed with the right hand using a dry-erase marker in which the ink was removed on an MR-compatible electronic drawing board (*Hybridmojo Touchpad)*. The drawing board operated like a traditional mouse pad, providing participants with real-time feedback on their motion dynamics. Participants were seated with the drawing board placed on their lap. Template shapes and tracing feedback were presented on a computer screen in front of the participant while their hands and the drawing board were covered. Participants' hands were occluded to obtain an experience similar to that of the scanning session (where their hands could not be seen). The behavioral session began with a familiarization phase, during which participants became acquainted with the task and setup. This was followed by a trace-collecting phase, during which traces were collected in order to present later in the trace-viewing condition within the MRI.

During the fMRI scanning session, participants were engaged in either tracing or viewing the same template shapes as in the behavioral session while their BOLD responses were recorded. Each participant completed three experimental conditions: active tracing, trace-viewing, and static template viewing (comprised of two experimental runs each). The active tracing condition was like that performed outside the scanner. Participants were engaged in a shape-tracing task on the same MR-compatible drawing board but were supine inside the scanner. In the trace-viewing condition, participants were instructed to observe a replay of the dynamic traces of the same two shape templates they produced before entering the scanner (see Figure 4A). During the static-viewing condition, participants were presented with the shape templates (without any traces) for the same amount of time as in the tracing and trace-viewing conditions (see Figure 4B). Templates and trace feedback were presented on an MR-compatible screen viewed from a head-mounted mirror, as commonly used in imaging experiments (see Figure 4C). The order of run types was counterbalanced across participants, and the order of presented traces was randomized (see Figure 4D).'

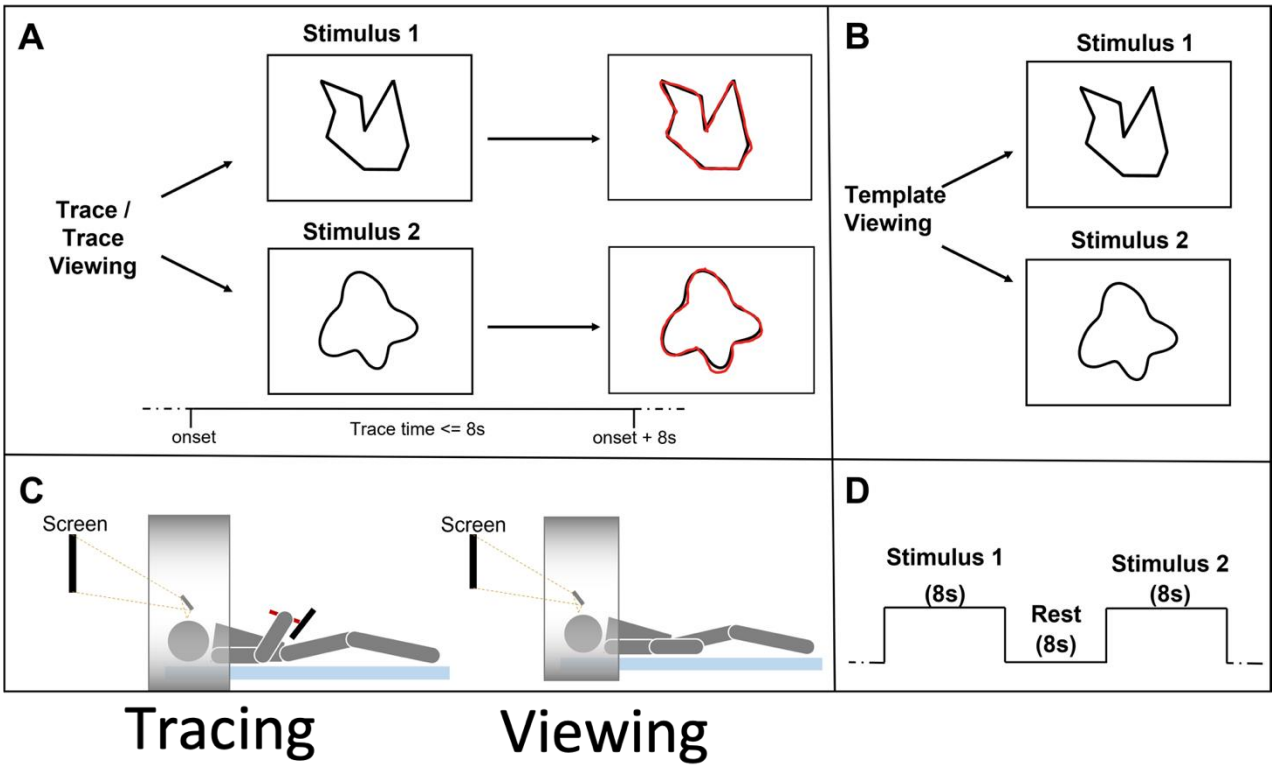


Figure 4: Experiment Design

A. Timeline of a single event in the active tracing and trace-viewing conditions. Participants were informed of the condition at the beginning of each run (either trace or observe previously collected traces). Template and traces were presented for 8s. B. Design of the template viewing condition. Templates were presented for 8s.C. Position of the participant in the scanner while tracing (left panel) and trace/template viewing (right panel). D. Schematic timeline of experiment run (similar for all conditions, one condition per run). The order of stimuli presentation was randomized during each run.

To keep participants attentive to the presented visual stimuli, we used an attention task, requiring them to detect changes in color of the template shape during all conditions. Each run consisted of either two or four catch trials. At the end of each run, participants had to report the number of trials in which they detected a color change. Functional data from catch trials were not included in the final analysis.
In addition to the experimental runs, this study contained functional localizers designed to locate visual brain regions sensitive to shapes and motion. During the shape localizer, participants were presented with the two shapes used in the experimental condition filled with a checkerboard pattern to increase neural activation in visual regions. Shapes were presented in a block design, each containing one flashing shape on the screen for eight seconds. For the motion localizer, participants were presented with an array of dots in a block design (Huk et al., 2002; Martins, 2018). The blocks alternated between the presentation of static dots and dots radiating to and from the center of the screen.

fMRI Data Acquisition

Functional neuroimaging was conducted at the Strauss Center for Computational Neuroimaging at Tel-Aviv University, using protocols similar to those applied in previous publications (Aberbach-Goodman et al., 2022; Buaron et al., 2020). In addition to functional scans, a whole-brain high-resolution T1-weighted scan was acquired for each participant for anatomical reference. Functional data pre-processing included brain extraction, slice-time correction, high-pass filtering at 100s (0.01Hz), motion correction to the middle time-point of each run, and autocorrelations correction. We excluded data from a condition's analysis when a participant's absolute

displacement values exceeded 3.5 mm for either of its runs. Eight participants were excluded due to technical problems, resulting in a sample of 14 participants.

Data Analysis

Kinematic Analysis: The kinematic data for both active tracing (traces completed inside the scanner) and passive trace-viewing (traces completed prior to entering the scanner) were analyzed to determine if differences in classification accuracy can be attributed to differences in kinematics inside and outside the scanner. Specifically, we computed the differences in measurements for each participant between the two shapes for both conditions. To determine if the differences between the shapes were consistent across conditions, a paired t-test was performed comparing the absolute value of the Δ between shape 1 and shape 2 of the active tracing condition with that of the passive trace-viewing condition. For this analysis, we did not perform a correction for multiple comparisons, the more stringent criteria could potentially lead us to disregard potential differences in this control. In our kinematic analysis, we investigated several derived measures, including maximum velocity, mean velocity, maximum acceleration, mean acceleration, total duration of movement (time elapsed from onset of movement to completion of the trace), total trace duration (time elapsed from presentation of shape to completion of shape), median velocity, SD velocity, median acceleration, and SD acceleration, smoothness (calculated using the Spectral Arc Length method), and trace accuracy. The accuracy was calculated as the percentage of pixels within the perimeter of each trace that matched between the trace and the template (see Figure 5).

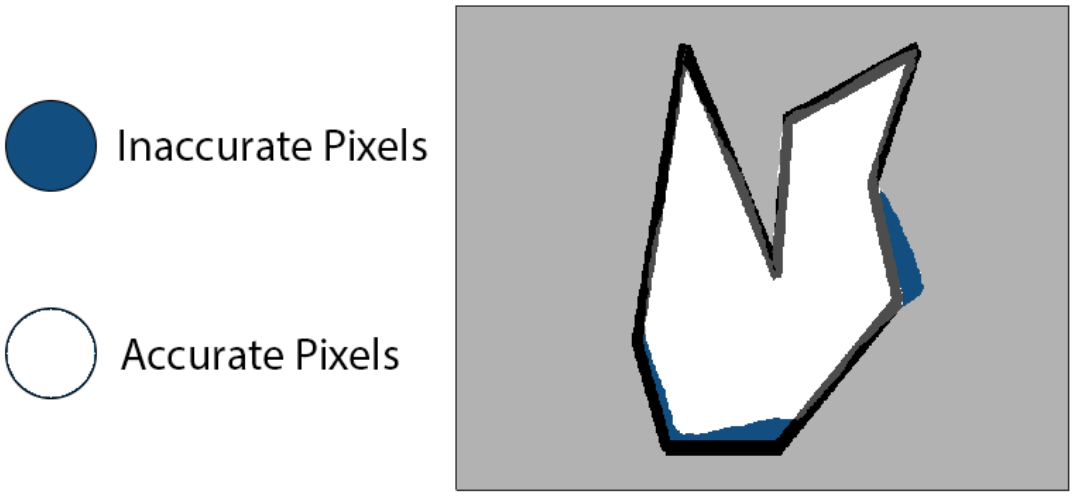


Figure 5: Sample trace demonstrating how trace accuracy was calculated. Blue pixels represent inaccurate pixels, white pixels represent accurate pixels, black pixels represent outline of the template shape. Accuracy was calculated as:

$$Trace\ Accuracy = (Total\ Pixels\ in\ template - Inaccurate\ Pixels) \Big/ Total\ Pixels\ in\ template$$

Localizer analysis: Data from both localizer runs were analyzed using a general linear model approach. In the shapes localizer, we used the contrast of [(shape 1 + shape 2) > rest] to detect visual areas sensitive to the presentation of our shapes (shape region of interest). In the motion localizer, we used the contrast of [moving dots > static dots] to detect areas sensitive to the motion of visual shapes (motion region of interest). We used these localizers to generate multi-subject, Montreal Neurological Institute (MNI)-normalized maps. We generated binary masks for each localizer representing significant voxels using the group-level maps for the motion and shape localizers ($p < 0.01$). A new binary mask was then calculated, presenting the union of these masks, indicating voxels below the threshold in either of the localizers. Lastly, the final region of interest (ROI) was generated by intersecting the union mask of the localizers with a mask of the voxels shared by all participants.

ROI Analysis: We generated two additional anatomic ROIs using the Harvard-Oxford cortical structural atlases for the inferior division of the lateral occipital cortex and the

occipital pole. Anatomic regions were defined as the intersection of voxels that were above 50% probability within each region and those within the localizer mask.

Multi-voxel pattern analysis (MVPA): MVPA was used to classify shape identity in different conditions to address the experimental questions. A support vector machine (SVM) classifier (Chang & Lin, 2011) was implemented to discriminate between activation patterns evoked by the two shapes in the ROIs defined by the localizers and the anatomical regions of interest. To this end, we conducted the following steps on the data of each participant separately.

For each voxel in the brain and each stimulus event, we calculated the average percent signal change across the first seven seconds of the block, offset by two seconds to account for the hemodynamic response function relative to the time course means. For each voxel, defined as center-voxel, we outlined a neighborhood that included the center voxel and its 26 closest voxels (in Euclidean distance). The signal pattern from this center voxel and neighboring voxels was extracted for each experimental trial. Using a leave-one-out approach, we trained an SVM classifier to discriminate signal patterns from trials of the two shapes and tested accuracy levels in discriminating left-out samples. We calculated the average accuracy level on the test set across all iterations of leave-one-out and assigned this as the decoding accuracy between shapes of the center voxel. After calculating the decoding accuracy of shapes for each voxel in the above manner, mean accuracy across all voxels within each ROI was taken as the dependent measure for further analysis. Each ROI was restricted to include the same number of voxels across participants by selecting voxels which fell in the intersection of the union mask of the localizers with a mask of the voxels shared by all participants.

Between Condition Comparison: We compared the mean shape classification accuracy of active tracing and passive trace-viewing using one-tailed paired-sample t-tests. These comparisons were completed in the ROI as defined by the union of the two functional localizers (henceforth referred to as localizer ROI), and the intersection of the Localizer ROI and the occipital pole (OP ROI), as well as the intersection of the Localizer ROI and the lateral occipital cortex (LOC ROI), resulting in three main regions of interest. Each of the three regions of interest was compared in the right

hemisphere, the left hemisphere, and both hemispheres together. There were nine comparisons in total. We used Bonferroni correction as we performed multiple comparisons, resulting in a statistical significance level of $\alpha = 0.05/9 \approx 0.0055$. The p-values for this comparison are reported after adjustment of their uncorrected p-values by multiplication of the number of comparisons.

Between Hemisphere Comparisons: For each condition, we compared the mean shape classification accuracy between the left and right hemispheres using a two-tailed paired t-test. These comparisons were completed separately for each condition in the Localizer ROI, OP ROI, and LOC ROI, resulting in 6 total comparisons. A Bonferroni correction for multiple comparisons resulted in a statistical significance level of $\alpha = 0.05/6 \approx 0.0083$. The p-values for this comparison are reported after adjustment of their uncorrected p-values by multiplication of the number of comparisons.

Results

Functional Localizer and Anatomical ROIs: A total of 21,089 voxels were significant ($p < 0.01$) across both hemispheres in the Localizer ROI (see Figure 6). For a detailed description of the number of voxels in each ROI refer to Table 1.

Table 1: Description of number of Voxels in each ROI

ROI	Hemisphere	Number of voxels
Localizer ROI	Both	21089
	Right	11442
	Left	9647
Occipital Pole	Both	1704
	Right	904
	Left	800
Lateral Occipital Cortex	Both	1753
	Right	841
	Left	912

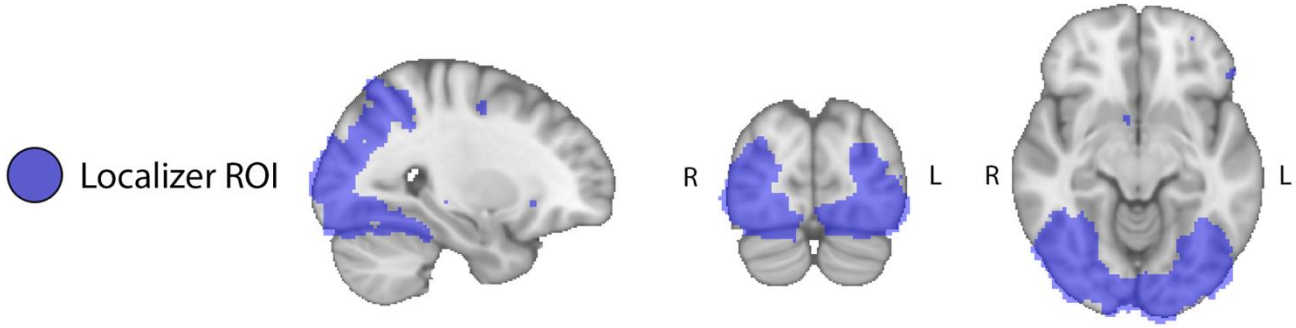


Figure 6: A Boolean map of the general linear model that results from the union of the two visual localizers (n=14, p<0.01). In the shape localizer, we contrasted [(shape 1 + shape 2) > rest] to detect visual areas sensitive to the presentation of our shapes (shape region of interest). In the motion localizer, we contrasted [moving dots > static dots] to detect areas sensitive to the motion of visual shapes This mask was defined in the MNI space and is presented on sagittal, coronal, and axial slices of the brain normalized to this space.

Between Condition Comparison: To assess whether interhemispheric differences in shape representation exist, and how voluntary action influences these differences, we compared mean classification accuracies between hemispheres independently for both the active tracing and passive trace-viewing conditions. In the bilateral Localizer ROI, the mean shape classification accuracy in the active tracing condition was higher than that of the passive trace-viewing. However, this difference was not statically significant following correction for multiple comparisons (active tracing: M = 0.562, SD = 0.037; trace-viewing: M = 0.535, SD = 0.037; t(13) = 2.13, p = .240 (adjusted), d = 0.57). This was also true when the left hemisphere (active tracing: M = 0.574, SD = 0.045; trace-viewing: M = 0.535, SD = 0.038; t(13) = 2.74, p = .076 (adjusted), d = 0.73), and right hemisphere (active tracing: M = 0.551, SD = 0.036; trace-viewing: M = 0.534, SD = 0.038; t(13) = 1.28, p = 1.000 (adjusted), d = 0.34) were examined separately (see Figure 7).

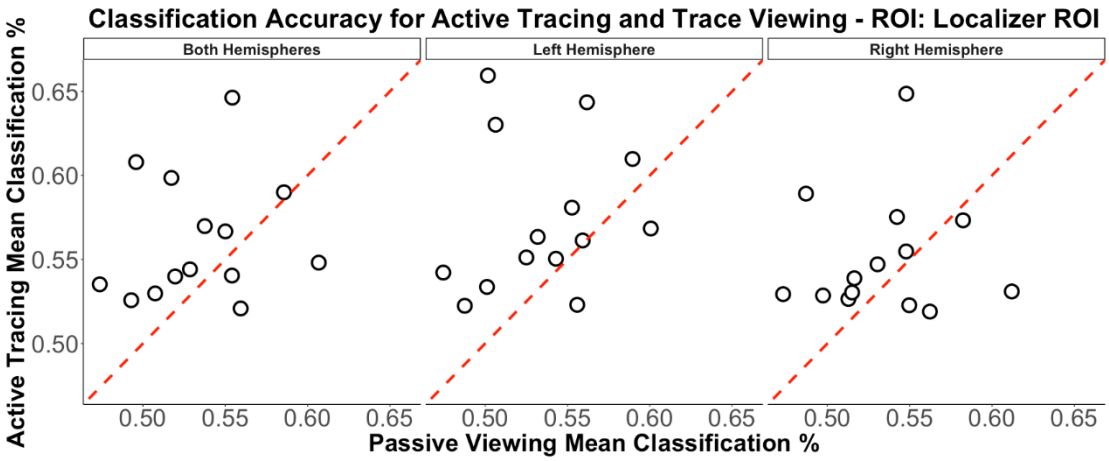


Figure 7: Scatter plot comparing each participant's mean shape classification accuracy in the active tracing and passive trace-viewing conditions across all voxels in the Localizer ROI. Each dot represents data from a single subject, and the red line represents identical decoding accuracies. Although most participants are above the diagonal (i.e. higher classification levels in the active condition), none of the comparisons were statistically significant after correction for multiple comparisons.

In the bilateral OP ROI, the mean classification accuracy in the active tracing condition was higher than that of the passive trace-viewing condition, and the difference was statistically significant (active tracing: M = 0.670, SD = 0.051; trace-viewing: M = 0.619, SD = 0.071; t(13) = 3.51, p = .017 (adjusted), d = 0.94). A similar pattern was found in the left hemisphere of the OP ROI, (active tracing: M = 0.662, SD = 0.057; trace-viewing: M = 0.600, SD = 0.069; t(13) = 3.70, p = .012 (adjusted), d = 0.99), as well as the right hemisphere of the OP ROI (active tracing: M = 0.678, SD = 0.059; trace-viewing: M = 0.640, SD = 0.078; t(13) = 2.47, p = .127 (adjusted), d = 0.66). However, for the right hemisphere the results failed to reach significance following correction for multiple comparisons (see Figure 8).

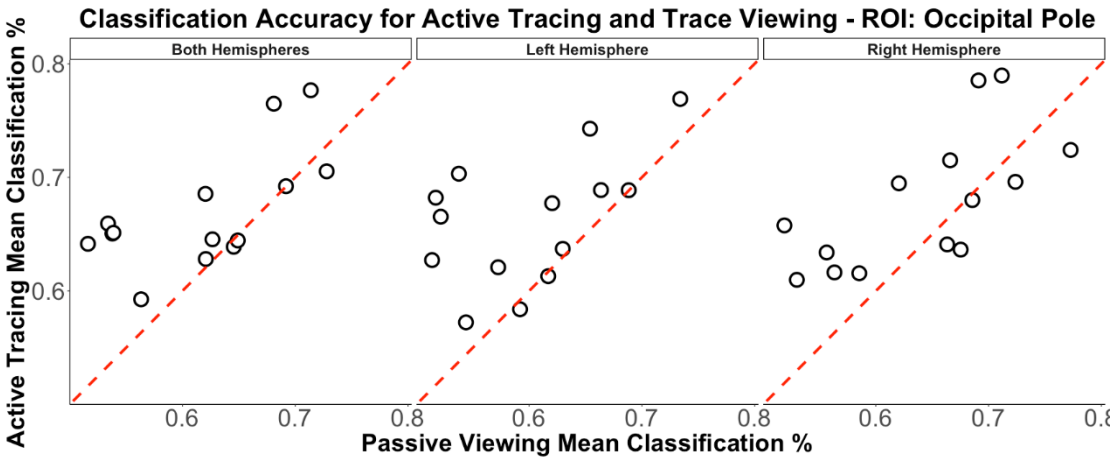


Figure 8: Scatter plot comparing each participant's mean shape classification accuracy in the active tracing and passive trace-viewing conditions in the OP ROI (conventions as in Figure 7). The differences in the left hemisphere and in both hemispheres together were statistically significant also after correction for multiple for comparisons.

In the LOC ROI, the mean shape classification accuracy in the active tracing condition was not significantly higher than that of the passive trace-viewing condition (active tracing: M = 0.543, SD = 0.047; trace-viewing: M = 0.517, SD = 0.042; t(13) = 1.67, p = 0.531 (adjusted), d = 0.45). Similar results were found in the left hemisphere of the LOC ROI (active tracing: M = 0.553, SD = 0.054; trace-viewing: M = 0.517, SD = 0.052; t(13) = 1.97, p = .320 (adjusted), d = 0.53), and in the right hemisphere of the LOC ROI (active tracing: M = 0.534, SD = 0.048; trace-viewing: M = 0.517, SD = 0.038;t(13) = 1.08, p = 1.000 (adjusted), d = 0.29) (see Figure 9).

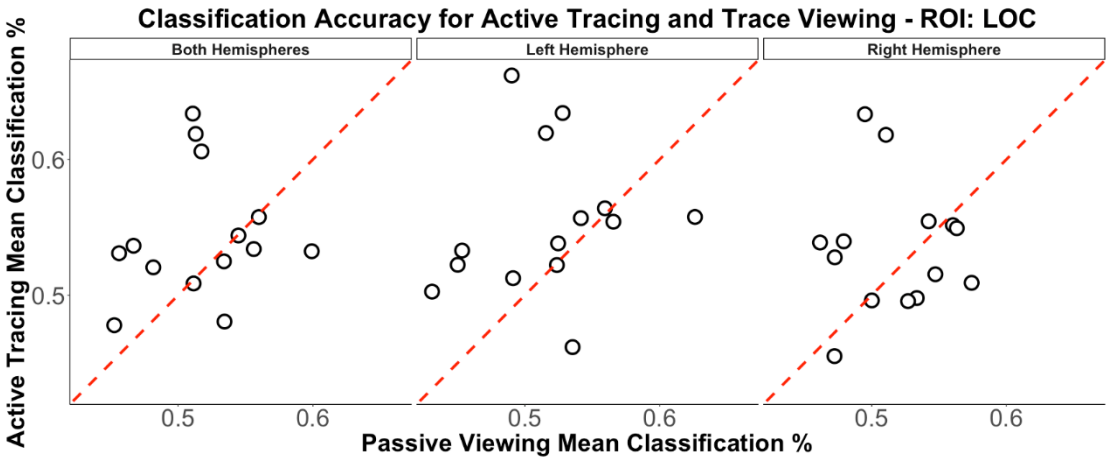


Figure 9: Scatter plot comparing each participant's mean shape classification

accuracy in the active tracing and passive trace-viewing conditions based on activity patterns in the LOC ROI (conventions as in Figure 7). None of the comparisons were statistically significant.

Figure 10 illustrates the mean shape classification accuracy for all regions of interest for both active and passive tracing conditions. In Figure 11, the shape classification accuracy across the two conditions is represented at the group level, highlighting the voxels that effectively discriminate between the shapes.

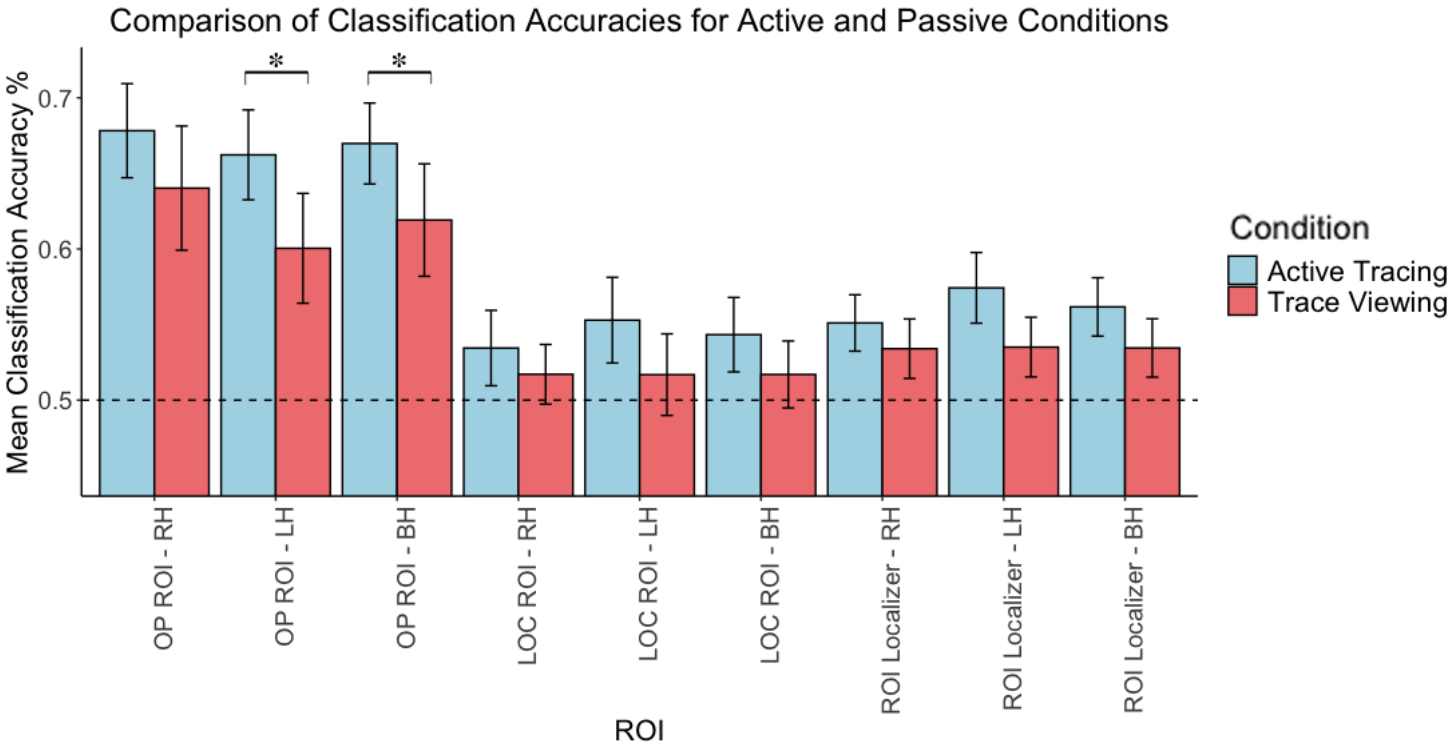


*Figure 10: The mean shape classification for the active tracing condition was higher than that of the passive trace-viewing condition for all regions of interest. In addition, mean shape classification accuracy values in the occipital pole were higher than those in the lateral occipital cortex, regardless of hemisphere. Acronyms used on the x-axis are structured as follows: the first part refers to the region (e.g., "OP" for occipital pole, "LOC" for lateral occipital cortex) while the suffix indicates the hemisphere ("RH" for right hemisphere, "LH" for left hemisphere, and "BH" for both hemispheres). Note: * indicates $p < .05$.*

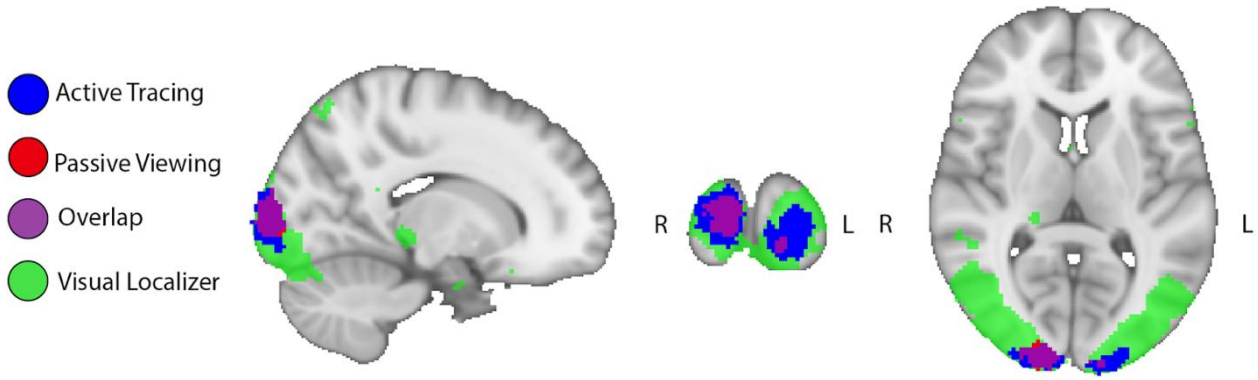


Figure 11: A Boolean map of center voxels differentiating between shapes 1 and 2 with classification accuracy greater than 55% across all participants. Blue voxels represent voxels with a mean classification accuracy above threshold in the active tracing condition, red voxels represent voxels with a mean classification accuracy above threshold in the passive trace-viewing condition, purple voxels represent voxels with a mean classification accuracy above threshold in both conditions. Green voxels represent the region of interest as defined by visual localizers.

Between Hemisphere Comparisons: In the active tracing condition, we compared the decoding accuracies across hemispheres but found that none of the comparisons were significant following correction for multiple comparisons (see Figure 12). For a detailed description of statistical tests for between hemisphere comparisons see Table 2.

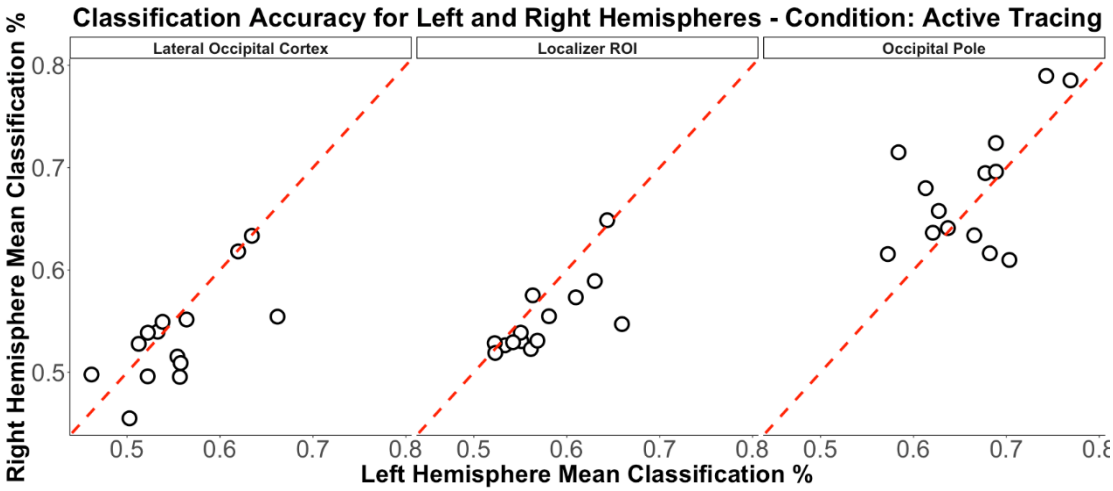


Figure 12: Scatter plot comparing each participant's mean shape classification accuracy in the active tracing condition between the right and left hemispheres for the three regions of interest (conventions as in Figure 7).. None of the comparisons were statistically significant after correction for multiple comparisons.

In the trace-viewing condition, there was no significant difference between the left and right hemispheres in the Localizer, as well as between the left and right hemispheres in the LOC ROI. However, when comparing the left and right hemispheres in the OP ROI, a significant difference in mean shape classification accuracies between the hemispheres was found (see Figure 13 right panel). For a detailed description of statistical tests between hemisphere comparison see Table 3.

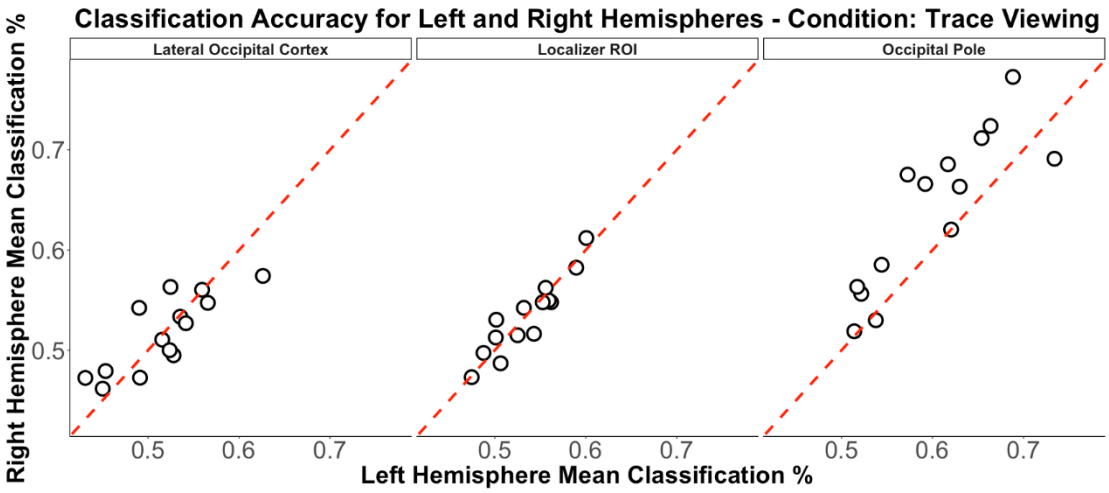


Figure 6: Scatter plot comparing each participant's mean shape classification accuracy in the passive trace-viewing condition between the right and left hemispheres for the three regions of interest (conventions as in Figure 7). The comparison in the occipital pole was statistically significant.

Table 2: Comparison of Mean Shape Classification Between Hemispheres for Active Tracing and Trace-Viewing Conditions

Condition	ROI	Hemisphere	Mean (M)	Standard Deviation (SD)	t(13)	Adjusted p-value	Effect Size (d)
Active Tracing	Localizer ROI	Left	0.574	0.045	2.79	0.092	0.75
		Right	0.551	0.036			
	OP ROI	Left	0.662	0.057	-1.09	1.000	0.29
		Right	0.678	0.059			
	LOC ROI	Left	0.553	0.054	1.78	0.587	-0.48
		Right	0.534	0.048			
Trace-Viewing	Localizer ROI	Left	0.535	0.038	0.26	1.000	-0.069
		Right	0.534	0.038			
	LOC ROI	Left	0.517	0.052	-0.03	1.000	0.008
		Right	0.517	0.038			
	OP ROI	Left	0.600	0.069	-3.73	**0.015***	0.998
		Right	0.640	0.078			

Note: Values with an asterisk (*) have a p-value of less than 0.05 following correction for multiple comparisons and are statistically significant.

Trace Kinematics: In addition to the neural signals, we measured trace kinematics. In principle, kinematic differences between shape 1 and shape 2 could result in different neural signals, leading to significant decoding differences between conditions. It is possible, however, that larger differences in kinematic measures between shapes in the active condition may explain the higher accuracy of decoding. To address this issue, we examined various shape kinematic differences between conditions. The difference in shape kinematics between conditions was not significant for mean acceleration, median acceleration, mean velocity, maximum velocity, median velocity, SD of velocity, smoothness, time to complete shape, trace accuracy, and trace duration from the onset of movement. The only two measures in which a difference was found were maximum acceleration and SD of acceleration. However, for both measures, the difference was greater in the passive trace-viewing condition

(see Figure 14) (for a detailed description of statistical tests for each kinematic metric see Table 3).

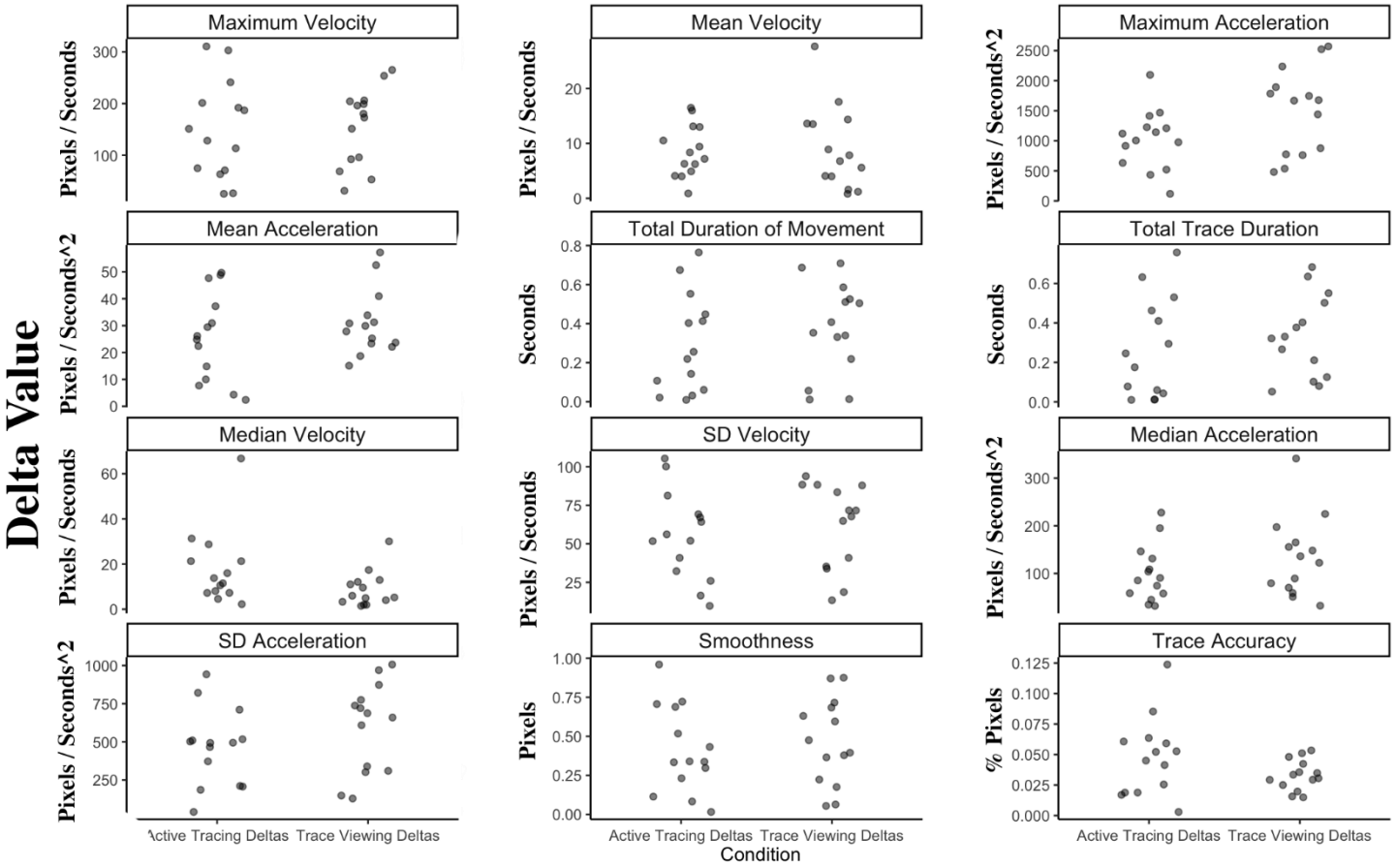


Figure 14: Scatterplot of the differences in each kinematic measure between shape 1 and shape 2 for each participant in each condition. Each point represents the difference between shape 1 and shape 2 for a single participant on the respective measure.

Table 3: Statistical tests for the kinematic metrics

Metric	Condition	M	SD	t(13)	p-
Acceleration Metrics $\left({}^{pixels}/_{seconds^2}\right)$					
Mean acceleration	Active	24.5	16.3	-1.370	0.19
	Trace-	30.9	12.1		
Maximum acceleration	Active	1019.8	495.2	-3.64	**<0.01***
	Trace-	1497.5	708.0		
Median acceleration	Active	99.3	58.8	1.95	0.07
	Trace-	133.7	82.6		
SD of acceleration	Active	462.5	251.1	-2.20	**<0.05***
	Trace-	590.5	293.1		
Velocity Metrics $\left({}^{pixels}/_{seconds}\right)$					
Mean velocity	Active	8.6	4.7	-0.24	0.81
	Trace-	9.1	7.5		
Maximum velocity	Active	149.1	94.2	-0.286	0.78
	Trace-	154.9	74.5		
Median velocity	Active	17.9	16.6	1.74	0.11
	Trace-	8.7	7.8		
SD of velocity	Active	55.2	28.8	-1.01	0.33
	Trace-	61.4	27.7		
Other Metrics					
Smoothness $(pixels)$	Active	0.41	0.28	-0.73	0.48
	Trace-	0.46	0.28		
Time to complete shape $(seconds)$	Active	0.27	0.25	-1.03	0.32
	Trace-	0.33	0.21		
Trace accuracy $\left(\frac{(Total\ Pixels\ in\ template - Inaccurate\ Pixels)}{Total\ Pixels\ in\ template}\right)$	Active	-0.047	0.032	-1.75	0.10
	Trace-viewing	-0.033	0.012		
Trace duration from onset of movement $(seconds)$	Active	0.29	0.25	-1.09	0.30
	Trace-	0.38	0.23		

Note: Values with an asterisk (*) have an uncorrected p-value of less than 0.05 and are statistically significant.

Figure 15 demonstrates all the traces of one participant overlaid on each other.

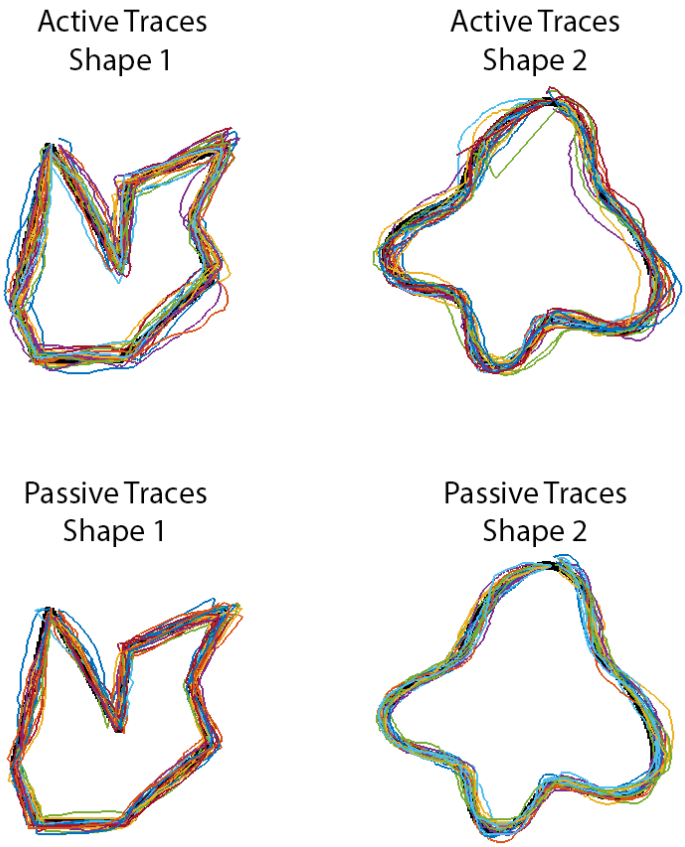


Figure 15: All traces of one participant (Participant 5) overlaid.

Behavior in Catch Trials: After each run, participants were instructed to report the number of catch trials they observed. None of the subjects incorrectly reported more than one catch trial in their fMRI session. Given that there were two runs for every condition, this culminated in a total of 28 opportunities to report catch trials for each condition across all participants. In the active tracing condition, participants correctly reported the number of catch trials in 25 out of 28 opportunities (89.3%), whereas in the trace-viewing condition they correctly reported the catch trials 26 out of 28 opportunities (92.9%) with no significant difference between conditions (X^2 (1, N = 56) = 0.22, p = 0.64).

Discussion

In this study, we investigated whether motor engagement sharpens the neural representation of visual shapes. A within-subject design was employed to explore this question. Participants completed two primary experimental conditions inside an MRI:

1) actively tracing two distinct shapes while receiving real-time feedback about their traces, 2) viewing the dynamic traces of the two shapes they created prior to entering the MRI. We hypothesized that motor engagement would sharpen the neural representation of visual shapes in the visual cortex. Specifically, we used multi-voxel pattern analysis to decode which of the two shapes the participant saw based on their BOLD activity. We predicted that within the visual cortex, the active tracing condition would result in higher shape classification accuracies compared to the trace-viewing condition. Our results supported this hypothesis.

We found that in the active tracing condition, participants had significantly higher mean shape classification accuracies (relative to the passive trace-viewing condition) in the bilateral OP ROI, as well as the left OP ROI. Furthermore, the right OP ROI, the bilateral localizer ROI, and the left localizer ROI also showed higher mean classification accuracies in the active tracing condition, however, these differences failed to reach significance following correction for multiple comparisons. In the active tracing condition, the mean classification was higher than in the passive viewing condition in all regions. We interpret these findings as evidence for sharpening of the neural representation of visual shapes. As far as we are aware, this is the first study to demonstrate neuronal sharpening in the context of self-generated input as opposed to sensory input derived from external sources.

According to an extensive body of literature, voluntary action modulates its sensory consequences at both the neural and behavioral levels. Additionally, studies have shown that the way in which visual information is perceived is influenced by prior expectations (Kok et al., 2012). Moreover, observation of visual information that is congruent to actions, sharpens the neural representation of the observed stimuli. This was demonstrated by higher decoding accuracies of actions based on activity patterns in early visual regions when actions of an avatar were congruent to participants' hand actions (as opposed to incongruent actions) (Yon et al., 2018). These findings support Bayesian accounts of perception, which claim that prior expectations can provide an advantage when making predictions about sensory input (Yuille & Kersten, 2006). Our findings support previous research on sensory modulation and the role of motor engagement in sharpening the neural representation of sensory information. Additionally, this study demonstrates the advantage of voluntary action on the brain's

capacity to process shapes, further solidifying our understanding of the forward model by suggesting that rather than serving as a mechanism that primarily cancels the redundant reafferent information, under certain circumstances, the efference copy may sharpen the neural signal, tuning it towards the expected sensory outcomes. This study expands on the literature by demonstrating that motor engagement sharpens the neural representation of visual information, relative to passively viewed stimuli. These results hint at the possible existence of additional information encoded in motor engagement beyond the prediction of the sensory consequences of an action, since in the trace-viewing condition, participants also had some capacity to predict the formation of the trace. Future research could investigate whether the efference copy, traditionally seen as a predictor of sensory consequences of actions, may also encode kinematic information.

This study has uncovered some exploratory findings that could shine light on the manner in which motor engagement affects neural responses in sensory areas. In line with previous research (Koivisto & Revonsuo, 2003), in the OP ROI, passive trace-viewing exhibited significantly greater mean shape classification accuracies in the right hemisphere than in the left hemisphere. However, no difference in mean shape classification accuracy was found between hemispheres in the OP ROI during the active tracing condition. As all participants were right-handed, we can interpret this finding in two ways. One explanation is that motor engagement sharpens neural activity in sensory regions ipsilateral to the active motor cortex more than it does in contralateral sensory regions. This results in a larger increase in mean shape classification accuracy relative to the passive trace-viewing condition in visual regions ipsilateral to the active motor cortex. A second possible explanation is that the right-hemisphere advantage in the passive viewing is not retained due to a ceiling effect reached in the (higher) classification levels in the active condition. Our findings are compatible with previous research performed by Buaron et al. (2020), who found that sensory modulation was stronger for visual information presented on the ipsilateral side (than on the contralateral side) to the active effector on the behavioral level. Additionally, research by Reznik et al. (2014) found that monaural hearing thresholds were lower for sounds triggered by the ipsilateral hand. These results provide evidence that the lateral relationship between an effector and the sense organ influences the manner in which the incoming sensory information is processed, and

that these modulations may be stronger in an ipsilateral configuration. Our findings provide neural evidence in support of this notion.

The lateral occipital cortex is conventionally considered to be the region with higher sensitivity towards shape (Amedi et al., 2007; Emberson et al., 2017). Nonetheless, in both the active and passive conditions, when comparing the mean shape classification accuracy in the occipital pole to that of the lateral occipital cortex, the occipital pole had significantly higher mean shape classification in all the comparisons (left hemisphere, right hemisphere, both hemispheres). This may be due to the fact that the traces only resulted in complete closed shapes toward the end of the block, and throughout most of the block, the information being processed was local features of the line such as line orientation, which is generally encoded in early visual regions (Lee, 2002).

Previous research has indicated that children who acquire knowledge of letters and numbers through the act of handwriting tend to perform better than those who rely solely on passive visual learning when it comes to recognizing symbols (Zemlock et al., 2018). Similarly, Seyll et al. (2020) found that handwriting resulted in better recognition of letter-like shapes than typing. In accordance with these findings, it has been demonstrated that the visual production of objects, for example through sketching, can reshape the objects' representational space (Fan et al., 2015). Our findings may provide an indication of the neural mechanisms underlying these behavioral findings.

The findings of this study have the potential to contribute to the clinical realm, specifically regarding rehabilitation. Individuals who have suffered from strokes often experience both perceptual and motor difficulties (Alt Murphy et al., 2017). If, as this study suggests, motor engagement generates a more distinct neuronal representation of sensory consequences, a rehabilitative strategy that combines perception with action may enhance recovery.

A possible alternative explanation to our results is that the higher classification in the active tracing condition could be attributed to increased attention to the shapes presented, rather than just the motor engagement involved. Research has shown that attention modulates both the perception and neural response to attended stimuli

relative to unattended stimuli (Hawkins et al., 1990; Maunsell & Treue, 2006). While this explanation is possible, to control attention we included “catch trials” in which participants were required to report the number of shapes they saw that were in a different color. Participants performed equally well in reporting catch trials across all conditions. Although we cannot entirely rule out attentional differences between active/viewing conditions, our behavioral results are less compatible with this alternative explanation.

The kinematic analysis revealed that the difference between two shapes differed between the active tracing and trace-viewing conditions in some of the measures. Specifically, when comparing maximum acceleration, and standard deviation of acceleration, participants exhibited significant differences between shape tracing kinematics inside and outside of the MRI. This presents a potential confounding variable as it is possible that the kinematic differences provided an additional source of information that enabled the classifier to correctly distinguish between the shapes. Given the higher classification accuracies in the active tracing condition, only larger differences in kinematics in the active condition would be compatible with better classification levels. Nonetheless, for both significant measures, the kinematic difference between the two shapes was actually greater for the passive trace-viewing condition. If indeed larger differences in kinematics provided more information to distinguish between the shapes, one would expect to find greater classification accuracy in the trace-viewing condition.

Another limitation of our study is that we cannot distinguish between the neural response modulations caused by motor engagement and those caused by tactile stimulation. Research by Bauer et al. (2009) found that a task-irrelevant tactile stimulus leads to modulation of both the neural and behavioral responses to visual stimuli. While we cannot disentangle motor from tactile engagement in our study, previous research conducted by Reznik, Ossmy, et al. (2015) attempted to do this by including a condition in which participants used a motion-sensitive glove that enabled the generation of sound by flexing a finger without making physical contact with a surface. This considerably minimized the amount of tactile feedback. The authors found that the signal difference in the left auditory cortex between the active and passive conditions remained consistent regardless of the degree of tactile feedback.

They concluded that in the auditory cortex, the modulation is driven by motor engagement rather than somatosensory feedback. A possible future study could compare the neural responses to sensory information generated by motor engagement, with the neural responses to sensory information generated by passive movement of the effector.

Conclusion

The results of this study indicate that motor engagement sharpens the neural representation of visual shapes in the visual cortex. Active tracing improved shape classification accuracy more than passive trace-viewing, particularly in the left and bilateral occipital poles. It also eliminated the discrepancy between the two hemispheres in terms of neuronal sharpening. This study yields a deeper understanding of the influence of action on perception and the neural mechanisms involved in this process and provides important insights towards shaping the forward model and the functional role of action on perception. These findings may also have implications for learning visual symbols and improving rehabilitation protocols.

References

Aberbach-Goodman, S., Buaron, B., Mudrik, L., & Mukamel, R. (2022). Same action, different meaning: neural substrates of action semantic meaning. *Cerebral Cortex*, *32*(19), 4293-4303 .

Ackerley, R., Hassan, E., Curran, A., Wessberg, J., Olausson, H., & McGlone, F. (2012). An fMRI study on cortical responses during active self-touch and passive touch from others. *Frontiers in behavioral neuroscience*, *6*, 51 .

Alt Murphy, M., Baniña, M. C., & Levin, M. F. (2017). Perceptuo-motor planning during functional reaching after stroke. *Experimental brain research*, *235*(11), 3295-3306 .

Amedi, A., Stern, W. M., Camprodon, J. A., Bermpohl, F., Merabet, L., Rotman, S., Hemond, C., Meijer, P., & Pascual-Leone, A. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature neuroscience*, *10*(6), 687-689 .

Bauer, M., Oostenveld, R., & Fries, P. (2009). Tactile stimulation accelerates behavioral responses to visual stimuli through enhancement of occipital gamma-band activity .*Vision research*, *49*(9), 931-942 .

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature neuroscience*, *1*(7), 635-640 .

Buaron, B., Reznik, D., Gilron, R. e., & Mukamel, R. (2020). Voluntary actions modulate perception and neural representation of action-consequences in a hand-dependent manner. *Cerebral Cortex*, *30*(12), 6097-6107 .

Butts, D. A., & Goldman, M. S. (2006). Tuning curves, neuronal variability, and sensory coding. *PLoS biology*, *4* ,(4)e92. https://journals.plos.org/plosbiology/article/file?id=10.1371/journal.pbio.0040092&type=printable

Cardoso-Leite, P., Mamassian, P., Schütz-Bosbach, S., & Waszak, F. (2010). A new look at sensory attenuation: Action-effect anticipation affects sensitivity, not response bias. *Psychological science*, *21*(12), 1740-1745 .

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, *2*(3), 1-27 .

Crapse, T. B., & Sommer ,M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews Neuroscience*, *9*(8), 587-600 .

Dewey, J. A., & Carr, T. H. (2013). Predictable and self-initiated visual motion is judged to be slower than computer generated motion. *Consciousness and cognition*, *22*(3), 987-995 .

Emberson, L. L., Crosswhite, S. L., Richards, J. E., & Aslin, R. N. (2017). The lateral occipital cortex is selective for object shape, not texture/color, at six months. *Journal of Neuroscience*, *37*(13), 3698-3703 .

Fan ,J. E., Yamins, D., & Turk-Browne, N. B. (2015). Common object representations for visual recognition and production. CogSci ,

Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., Faustman, W. O., & Roth, W. T. (2001). Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *American Journal of Psychiatry*, *158*(12), 2069-2071 .

Fukutomi, M., & Carlson, B. A. (2020). A history of corollary discharge: contributions of mormyrid weakly electric fish. *Frontiers in integrative neuroscience* .42 ,14 ,

Gibson, J. J. (1977). The theory of affordances. *Hilldale, USA*, *1*(2), 67-82 .

Gibson, J. J. (2014). *The ecological approach to visual perception: classic edition*. Psychology press .

Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurgery Clinics*, *22*(2), 133-139 .

Gritsenko, V., Krouchev, N. I., & Kalaska, J. F. (2007). Afferent input, efference copy, signal noise, and biases in perception of joint angle during active versus passive elbow movements. *Journal of Neurophysiology*, *98*(3), 1140-1154 .

Haggard, P., & Whitford, B. (2004). Supplementary motor area provides an efferent signal for sensory suppression. *Cognitive Brain Research*, *19*(1), 52-58 .

Hawkins, H. L., Hillyard, S. A., Luck, S. J., Mouloua, M., Downing, C .J., & Woodward, D. P. (1990). Visual attention modulates signal detectability. *Journal of experimental psychology: Human perception and performance*, *16*(4), 802 .

Helmholtz, H. L. F. v., & Southall, J. P. C. (1925). Treatise on physiological optics .

Hering, E. (1964). Outlines of a Theory of the Light Sense .

Hughes, G., & Waszak, F. (2011). ERP correlates of action effect prediction and visual sensory attenuation in voluntary action. *Neuroimage*, *56*(3), 1632-1640 .

Huk, A. C., Dougherty, R. F., & Heeger ,D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, *22*(16), 7195-7205. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6757870/pdf/ns1602007195.pdf

Isaacson, J. S., & Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron*, *72*(2), 231-243 .

Jia, H., Rochefort, N. L., Chen, X., & Konnerth, A. (2010). Dendritic organization of sensory input to cortical neurons in vivo. *Nature*, *464*(7293), 1307-1312 .

Kavroulakis, E., van Kemenade, B. M., Arikan, B. E., Kircher, T., & Straube, B. (2022). The effect of self-generated versus externally generated actions on timing, duration, and amplitude of blood oxygen level dependent response for visual feedback processing. *Human Brain Mapping*, *43*(16), 4954-496 .9

Koivisto, M., & Revonsuo, A. (2003). Object recognition in the cerebral hemispheres as revealed by visual field experiments. *Laterality: Asymmetries of Body, Brain and Cognition*, *8*(2), 135-153 .

Kok, P., Jehee, J. F., & De Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265-270 .

Laufer, Y., Hocherman, S., & Dickstein, R. (2001). Accuracy of reproducing hand position when using active compared with passive movement. *Physiotherapy research international*, *6*(2), 65-75 .

Lee, T. S. (2002). Top-down influence in early visual processing: a Bayesian perspective. *Physiology & behavior*, *77*(4-5), 645-650 .

Ling, S., Jehee, J. F., & Pestilli, F. (2015). A review of the mechanisms by which attentional feedback shapes visual selectivity. *Brain Structure and Function*, *220*, 1237-1250 .

Lubinus, C., Einhäuser, W., Schiller, F., Kircher, T., Straube, B., & van Kemenade, B. M. (2022). Action-based predictions affect visual perception, neural processing, and pupil size, regardless of temporal predictability. *Neuroimage*, *263*, 119601 .

Martins, R. (2018). *fMRI Localizer MT MST MT proper - MATLAB Psychophysics Toolbox*. In https://github.com/rmartins-net/Localizer-MT-MST-MT-proper-Psychophysics-Toolbox

Massaro, D. W., & Anderson, N. H. (1971). Judgmental model of the Ebbinghaus illusion. *Journal of experimental psychology*, *89*(1), 147 .

Maunsell, J. H., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in neurosciences*, *29*(6), 317-322 .

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural networks*, *9*(8), 1265-1279 .

Mifsud, N. G., Oestreich, L. K., Jack, B. N., Ford, J. M., Roach, B. J., Mathalon, D. H., & Whitford, T. J. (2016). Self-initiated actions result in suppressed auditory but amplified visual evoked components in healthy participants. *Psychophysiology*, *53*(5), 723-732 .

Ody, E., Straube, B., He, Y., & Kircher, T. (2023). Perception of self-generated and externally-generated visual stimuli :Evidence from EEG and behavior. *Psychophysiology*, e14295 .

Reznik, D., Guttman, N., Buaron, B., Zion-Golumbic, E., & Mukamel, R. (2021). Action-locked neural responses in auditory cortex to self-generated sounds. *Cerebral Cortex*, *31*(12), 5560-5569 .

Reznik, D., Henkin, Y., Levy, O., & Mukamel, R. (2015). Perceived loudness of self-generated sounds is differentially modified by expected sound intensity. *PloS one*, *10*(5), e0127651 .

Reznik, D., Henkin, Y., Schadel, N., & Mukamel, R. (2014). Lateralized enhancement of auditory cortex activity and increased sensitivity to self-generated sounds. *Nature communications*, *5*(1), 4059 .

Reznik, D., & Mukamel, R. (2019). Motor output, neural states and auditory perception. *Neuroscience & Biobehavioral Reviews*, *96*, 11 .6-126

Reznik, D., Ossmy, O., & Mukamel, R. (2015). Enhanced auditory evoked activity to self-generated sounds is mediated by primary and supplementary motor cortices. *Journal of Neuroscience*, *35*(5), 2173-2180 .

Reznik, D., Simon, S., & Mukamel, R. (2018 .(Predicted sensory consequences of voluntary actions modulate amplitude of preceding readiness potentials. *Neuropsychologia*, *119*, 302-307 .

Salzman, C. D., Britten, K. H., & Newsome, W. T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature*, *346*(6280), 174-177 .

Schafer, E. W., & Marcus, M. M. (1973). Self-stimulation alters human sensory brain responses. *Science*, *181*(4095), 175-177 .

Seriès, P., Latham, P. E., & Pouget, A. (2004). Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nature neuroscience*, *7*(10), 1129-1135 .

Seyll, L., Wyckmans, F., & Content, A. (2020). The impact of graphic motor programs and detailed visual analysis on letter-like shape recognition. *Cognition*, *205*, 104443 .

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2005). Evidence for sensory prediction deficits in schizophrenia. *American Journal of Psychiatry*, *162*(12), 2384-2386 .

Shergill, S. S., White, T. P ,.Joyce, D. W., Bays, P. M., Wolpert, D. M., & Frith, C. D. (2014). Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. *JAMA psychiatry*, *71*(1), 28-35 .

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of comparative and physiological psychology*, *43*(6), 482 .

Subramaniam, K., Kothare, H., Mizuiri, D., Nagarajan, S. S., & Houde, J. F. (2018). Reality monitoring and feedback control of speech production are related through self-agency. *Frontiers in human neuroscience*, *12*, 82 .

Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*(11), 745-756 .

Von Holst, E. (1954 .(Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behaviour* .

Voss, M., Ingram, J. N., Haggard, P., & Wolpert, D. M. (2006). Sensorimotor attenuation by central motor command signals in the absence of movement. *Nature neuroscience*, *9*(1), 26-27 .

Welniarz, Q., Worbe, Y., & Gallea, C. (2021). The forward model: a unifying theory for the role of the cerebellum in motor control and sense of agency. *Frontiers in Systems Neuroscience*, *15*, 644059 .

Witt, J. K ,.Proffitt, D. R., & Epstein, W. (2004). Perceiving distance: A role of effort and intent. *Perception*, *33*(5), 577-590 .

Witt, J. K., Proffitt, D. R., & Epstein, W. (2005). Tool use affects perceived distance, but only when you intend to use it. *Journal of experimental psychology: Human perception and performance*, *31*(5), 880 .

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature neuroscience*, *3*(11), 1212-1217 .

Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C .(2018) .Action sharpens sensory representations of expected outcomes. *Nature communications*, *9*(1), 1-8 .

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences*, *10*(7), 301-308 .

Zemlock, D., Vinci-Booher, S., & James, K. H. (2018). Visual–motor symbol production facilitates letter recognition in young children. *Reading and Writing*, *31*, 1255-1271 .