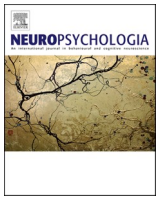
Neuropsychologia 221 (2026) 109329

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia



Neural correlates of 3D versus 2D perception: An activation likelihood estimation meta-analysis

Luiza P. Volpi [a], R. Nathan Spreng [b,c,d,e], Reza Farivar [a,*]

[a] Department of Ophthalmology and Visual Sciences, McGill University, Department of Neurology, Montreal, Quebec, Canada
[b] Montreal Neurological Institute, Department of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada
[c] McConnell Brain Imaging Centre, McGill University, Montreal, Quebec, Canada
[d] Departments of Psychiatry and Psychology, McGill University, Montreal, Quebec, Canada
[e] Douglas Mental Health University Institute, Verdun, Quebec, Canada

ARTICLE INFO

Keywords:
fMRI
3D shape
Depth cue
Dorsal stream
Ventral stream
Meta-analysis

ABSTRACT

Despite hundreds of neuroimaging studies examining the neural correlates of 3D shape perception (as opposed to 2D), there is no consensus because of the diversity of stimuli and depth cues used. We addressed this problem through an activation likelihood estimation (ALE) coordinate-based meta-analysis, pooling together studies that examined the 3D vs 2D shape contrast across multiple depth cues used to render the 3D shapes. A systematic review was performed using Medline, PsychInfo and Embase databases and yielded 25 empirical studies after screening. Articles were split into cue types—disparity (11), motion (10), mental rotation (1), shading (2) and texture (2). We performed three sets of ALE-based coordinate analysis—full-sample ALE analysis, sub-analyses testing individual depth cues separately, and a contrast analysis between disparity-defined 3D shapes and monocularly-defined 3D shapes. Results for the full-sample analysis showed that 3D shape perception is widespread throughout the high-level visual cortex regardless of depth cue. Although cue-specific analyses were underpowered, some trends were observed. Disparity-defined 3D shapes seem to engage higher-level dorsal stream areas, including bilateral intraparietal sulcus (IPS). Motion-defined 3D shape recruited ventral stream regions associated with object recognition processes. Monocularly-defined 3D shapes recruited ventral stream areas, mainly the bilateral inferior lateral cortex and dorsal stream IPS for the right hemisphere. W et al.hen contrasted with disparity, monocularly-defined 3D shapes recruited the left lateral occipital cortex. The results suggest laterality in 3D versus 2D shape representations and that 3D shape representations occur in both ventral and dorsal pathways regardless of the depth cues that define them.

Key points

- fMRI meta-analysis of 3D shape perception when contrasted with their 2D controls.
- High-level dorsal and ventral streams represent 3D shapes regardless of depth cue.
- Disparity-defined 3D shape representations are stronger in bilateral dorsal stream whereas monocularly-defined 3D shape representations are stronger in left ventral stream, suggesting that segregation and lateralization of reliable activation is dependent on depth cue.

1. Introduction

Real objects are processed differently than 2D images of those same objects (Chainay et al., 2010; Snow et al., 2011) meaning that 3-dimensionality is a relevant feature of object recognition, but neuroimaging studies of object recognition have largely used 2D images or 2D shapes (Andresen et al., 2009; Bernstein et al., 2014; Bracci and Op de Beeck,

* Corresponding author.
E-mail address: reza.farivar@mcgill.ca (R. Farivar).

https://doi.org/10.1016/j.neuropsychologia.2025.109329
Received 29 April 2025; Received in revised form 20 November 2025; Accepted 21 November 2025
Available online 24 November 2025

2016; Bracci et al., 2019; Konkle and Oliva, 2012; Mocz et al., 2023). Object recognition theories and models (Bülthoff and Edelman, 1992; Peissig and Tarr, 2006; Poggio and Riesenhuber, 1999; Riesenhuber and Poggio, 2000) have historically been focused on the hierarchical encoding of 2D features, leading to 2D shape recognition by using 2D—or edge-defined—features of object images (Bracci and Op de Beeck, 2016; Long et al., 2018; Ullman et al., 2016).

Three-dimensional features of objects are explicitly encoded in the primate brain, specifically in ventral stream areas traditionally associated with the processing of 2D features or simple shapes. Encoding of 3-dimensional oriented slants are separable from encoding 2-D edges or curvatures in the macaque V4 (Hinkle and Connor, 2002). Macaque V4 is also organized in subclusters of neurons that respond preferentially either to solid 3D shapes or to 2D patterns (Srinath et al., 2021). This segregation was observed by using 2-photon microscopy in V4 neurons while showing flat (i.e. shadow outline) versus solid shapes (i.e. shaded shape). Srinath et al. (2021) suggest that these clusters might serve different functions, such as extracting 3D surfaces of objects or 2D patterns that aid in overall recognition. Medial axis structure—or the core 'skeleton'—of objects is specifically encoded in the inferotemporal (IT) cortex of monkeys and, in humans, represented in V3 and lateral occipital complex (LOC; Lescroart and Biederman, 2013). In humans, scene-selective regions, such as the parahipoccampal place area and occipital place area also represent 3D structure of scenes (Lescroart and Gallant, 2019), although it is unclear whether this information is extracted with the purpose of 3D object perception or for facilitating interactions with the environment. Therefore, despite the success of 2D edge-based hierarchical models of ventral pathway processing of objects, there is extensive evidence that 3D information is an integral part of object representation.

In contrast to the central role assigned to ventral stream processing of objects, the dorsal stream's role in object recognition has been traditionally downsized, stemming from early lesion studies in the macaque (Mishkin et al., 1983) that were later replicated in humans (Goodale and Milner, 1992; Milner et al., 1991; Sereno et al., 2002). These studies showed a double dissociation between object discrimination (i.e., object recognition) and landmark detection (i.e., relationship between objects in a viewed scene) depending on the site of the lesion—ventral stream lesions in IT cortex affected object discrimination while dorsal stream inferior parietal cortex lesions affected landmark discrimination (Goodale et al., 1992; Milner et al., 1991; Mishkin et al., 1983; Sereno et al., 2002). Since then, the functional dichotomy of "what" and "where" has largely been attributed to the ventral and dorsal streams, respectively.

However, 3D depth information of objects is extensively processed in the dorsal stream. Studies by Dövencioğlu et al. (2013) and Murphy et al. (2013) suggest that integration of disparity with both shading and texture cues occur in the kinetic occipital/V3B—assigned to the dorsal stream—for simple 3D surfaces. In a ventral stream lesion study, Freud et al. (2017) showed that visual object agnosia patients had residual sensitivity for 3D structural information of abstract objects and that 3D object representations were generated in the dorsal stream as measured by fMRI, even though the ventral stream representations were largely compromised. Jeong and Xu (2016) showed that the human intraparietal sulcus (IPS) represents face identity of celebrities regardless of extensive image transformations.

While inferring 3D representations from 2D images (e.g., photographs, line drawings, etc.) can be ambiguous, the same cannot be said for representations inferred from individual depth cues. For example, stimuli rendered solely by texture gradients, disparity gradients, or velocity gradients (i.e., 3D structure-from-motion; SFM) do not have any useful edge information and highly capable 2D neural network models may not even detect an object from such cues, while humans not only can see them (Akhavein et al., 2018; Dehmoobadsharifabadi and Farivar, 2016; Farivar, 2009) but also carry out complex tasks, such as unfamiliar face recognition from pure gradient cues (Farivar et al., 2009; Liu et al., 2005). This can be easily appreciated with texture-defined 3D stimuli, because the edges of the texture can vary depending on the texture, but the gradient of the texture is the defining feature of the underlying 3D surface (see Fig. 1). Thus, objects defined by depth cues provide a unique probe of actual 3D representations, but as we see below, due to the diversity of stimuli, tasks, and depth cues, it has been difficult to define the cortical networks that contribute to processing of 3D object information.

Individual studies that assess the role of 3D shape processing as opposed to 2D shape are in general underpowered—each study recruited on average only 16 participants—and they not only vary by the type of depth cue used to render the 3D representations, but also in the types of stimuli and tasks performed. When in the scanner, participants can be asked to simply look at objects (Gillebert et al., 2015), perform an n-back task (Cavina-Pratesi et al., 2010), acuity judgements (Georgieva et al., 2009) and more. Also, some studies involved a psychophysical component which may have been related to stimulus learning (Georgieva et al., 2009; Georgieva et al., 2008; Gillebert et al., 2015) while others do not (Akhavein, 2017). It is worth noting that in the above-cited studies, the complexity of the stimuli used vary from simple shapes to complex objects including unfamiliar faces.

By pooling together—in a meta-analysis—studies that tested 3D shape representations over 2D ones regardless of their depth cue, stimuli or task, we can answer fundamental questions on the organization of 3D visual object recognition. Specifically, we ought to test (1) the ventral stream's role in object representation from 3D depth cues and (2) the dorsal stream's role in depth cue processing, integration, and complex shape representations.

1.1. Definitions

In the present study we consider "3D shapes" as ranging from simple to complex 3D objects, that can be either familiar or unfamiliar and by "2D shape" their 2D controls which are either 2D shapes or their depth cue-specific controls. We also use the terms **binocularly-defined 3D shapes and disparity-defines 3D shapes interchangeably.**

1.2. Aims

The present study aimed to first identify brain regions that are reliably engaged in 3D shape perception, regardless of depth cue, when contrasted with 2D shapes. Here, studies using a diversity of depth cues—such as gradients of shading, texture, binocular disparity, and motion velocity—to define the 3D shape percept have been selected. By selecting studies that contain high variability in lower-level features—i.e., the distinct depth cues—we can capture brain regions that reliably represent 3D shape as opposed to sensitivity to specific depth cues. In other words, given the heterogeneity of the studies (e.g., distinct depth cues, stimuli), brain areas that "survive" are therefore representing 3D structures or surfaces of objects when contrasted to their 2D shape controls. Using coordinate-based activation likelihood estimation (ALE) with the generalized 3D > 2D contrasts, we identified brain regions that were reliably implicated in 3D shape perception as opposed to 2D shape perception, and we were able to identify regions that converge in representing 3D shape regardless of depth cue. In this sense, we aimed to assess commonalities, or large-scale regions or networks that are overall representing 3D structures of objects when contrasted to their 2D counterparts, not differentiating the intricacies or specific depth cue processing or discriminating between them.

Second, because of the intrinsic variability of the contrasts chosen, we performed ALE-based sub-analyses for each type of depth cue. This was performed to distinguish between 3D shape-related activations that were specific to a unique depth cue and those that were cue-invariant. We wanted to disentangle depth cue information from that of 3D shape.

Lastly, evidence shows that binocular disparity is processed in widespread ventral and dorsal visual cortices (Abed Rabbo, Koch,

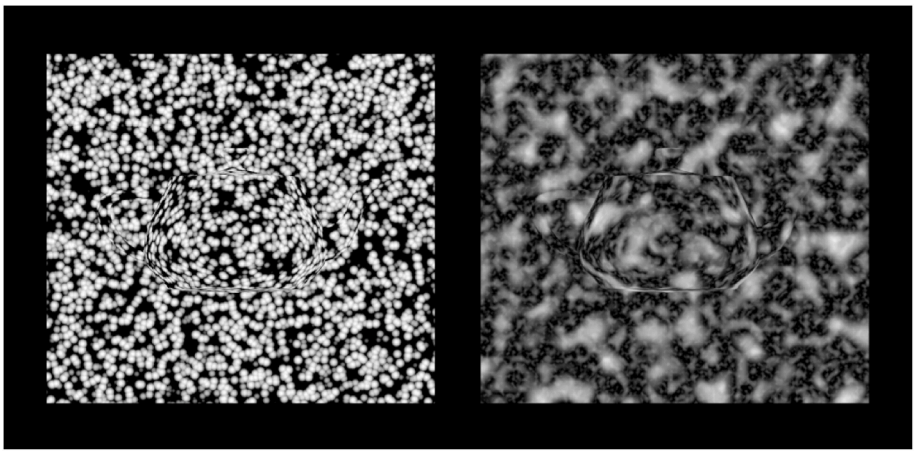


Fig. 1. A teapot defined by two distinct texture gradient cues.

Lefèvre and Seizeur, 2018; Tsao et al., 2003; Verhoef et al., 2016). To address which regions were reliably representing 3D shape information independently of 3D binocular disparity processing, we performed a third set of analyses, which contrasted monocularly-defined 3D shape percepts with disparity-defined 3D shape percepts (monocularly-defined 3D shapes > disparity-defined 3D shapes contrast). This contrast should instantiate regions that are fundamentally processing 3D structures or representing 3D shape information, and not only processing 3D or binocular vision. The complementary contrast—disparity-defined 3D shapes > monocularly-defined 3D shapes was also computed, for identification of regions that show stronger activation for disparity when the 3D shape information aspect is taken away from the equation.

2. Materials and methods

2.1. Literature search and study selection

We performed a systematic literature search with the aid of the Neuroscience Head Librarian (McGill University). The first step was setting up the research question in a way that pinpointed the main concepts for the search. This led to the extraction of four main concepts, here shown in bold, that were retrieved from the question: "What are the **neural correlates** of **3D shape perception** as opposed to 2D shapes, as defined by distinct **depth cues** in **humans**?". We used Ovid to search Medline, PsychINFO, and Embase databases and all retrieved articles, published until 14th of September (2021), were included in this meta-analysis. The search terms included both indexed search terms—which are unique for each database—and keywords—which were searched equally in all three databases, in the titles, abstracts and author-provided keywords domains. Table 1 provides detailed information on which keywords and specialized indexed search terms were used for the search. The indexed search terms were identified from a pre-identified set of articles (i.e., articles we were aware of and that we expected to find in the search; Fig. 2).

The search yielded 2807 articles, totalling 2391 after deduplication (Fig. 2). Deduping was performed in EndNote following the guidelines from Bramer et al. (2016). Studies were imported to Rayyan for title and abstract screening (https://www.rayyan.ai/, a free online tool for collaborative systematic review screening). Studies were included in the meta-analysis if they met the following criteria.

1. They employed fMRI or PET.
2. Were empirical articles (i.e., not reviews, theses, or meeting abstracts)
3. Tested 3D shape perception as contrasted with 2D shape perception in their reported group main effects. Here, "3D shape perception" entails simple and complex objects from which a 3D structure could be perceived because of the depth cue that rendered them (e.g., motion that constructs a 3-Dimensional percept). 2D shapes were matched controls in which the percept was a 2-Dimensional object in nature. For example, if a study contrasted structure-from-motion stimuli with static points, this contrast was not included as it would have needed to be contrasted with motion stimuli that revealed a 2D shape percept instead.
4. Used at least one type of depth cue to define the 3D shape percepts.
5. Used whole-brain coverage during fMRI or PET scan acquisitions.

Table 1
Terms used for the systematic database search.

		Headings/Subheadings Search Tags		
	Keywords*	MEDLINE	Embase	PsychINFO
Concept 1: fMRI/ PET	(fMRI or 'functional magnetic resonance imaging' or 'brain imaging' or PET or 'positron emission tomography').mp.	exp Magnetic Resonance Imaging/	exp functional magnetic resonance imaging/	exp Functional Magnetic Resonance Imaging/
Concept 2: 3D shape/structure recognition	('3D shape*' OR '3-D shape*' OR 'dimensional shape*' OR '2D shape*' OR '2-D shape*' OR surface* OR 'object recognition' OR 'object identi*').mp.	exp Depth Perception/or exp Form Perception/	exp Depth Perception/	exp "Form and Shape Perception"/or exp object recognition/or exp depth perception/
Concept 3: depth cues	('depth cue*' OR 'depth-cue*' OR shading* OR texture* OR 'binocular vision' OR 'binocular disparit*' OR stereo* OR 'random dot stereogram' OR SFM OR 'structure-from-motion' OR 'structure from motion').mp.	exp Cues/OR exp Vision, Binocular/OR exp Vision, Monocular/OR exp Vision Disparity/	exp binocular vision/or exp stereoscopic vision/or exp monocular vision/	exp Cues/or exp Binocular Vision/or exp Stereoscopic Vision/or exp Monocular Vision/
Concept 4: human	human*.mp.	exp Humans/	exp human/	n/a

Table Note: The same keywords were searched in all databases, in titles, abstracts and author-provided keywords. Indexed search terms are specific for each database and are reported under their respective column in the table.

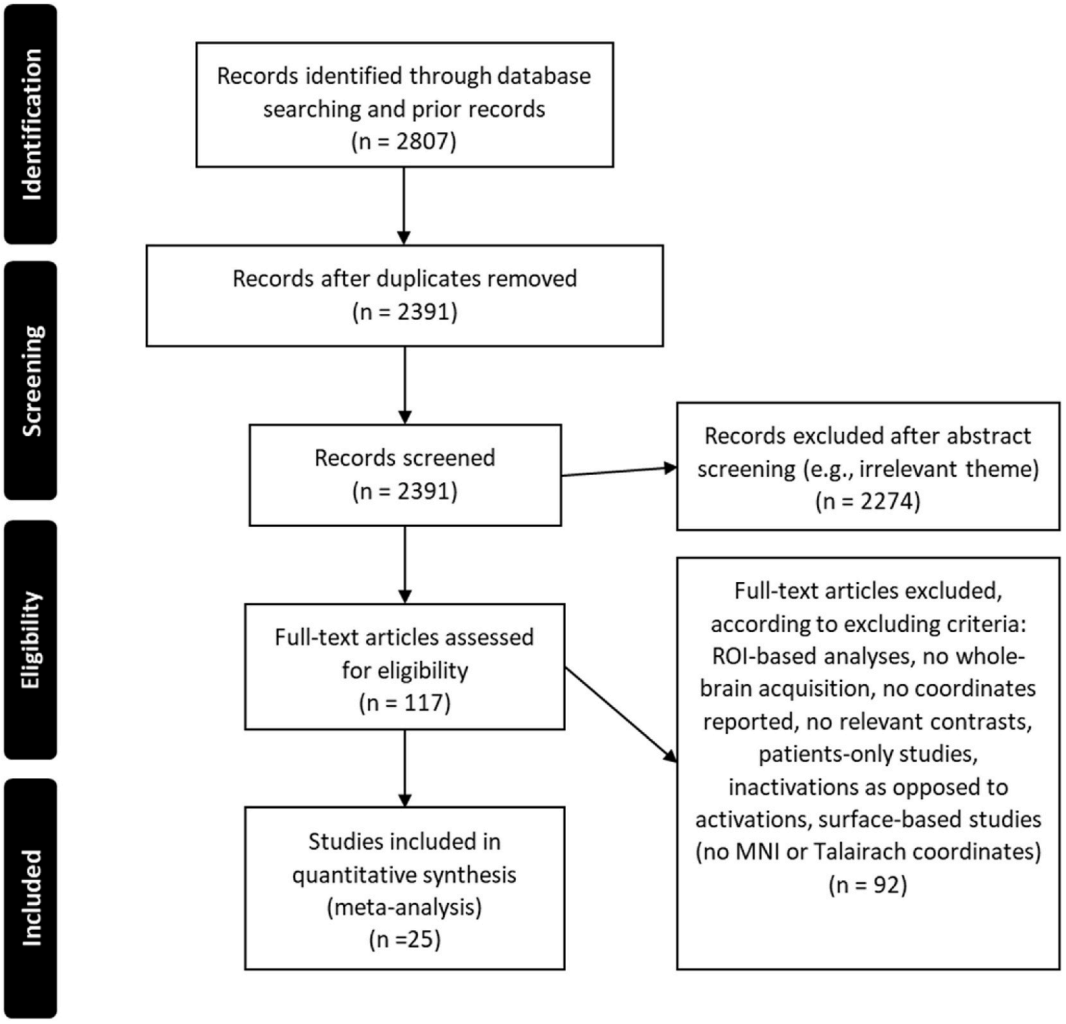


Fig. 2. Flowchart describing article selection, exclusion criteria and screening process, according to PRISMA guidelines (Moher et al., 2009, www.prisma-statement.org).

6. Performed whole-brain analyses (i.e., to avoid overestimation and bias of activation of a few areas over others, so papers that performed only ROI-based analyses or partial volume acquisition were excluded).
7. Reported either Montreal Neurologic Institute (MNI) or Talairach coordinates for peak activity.
8. Showed only activation—i.e. not deactivation—for 3D shape perception when contrasted to 2D shapes.
9. Studied healthy human subjects (i.e., had at least a healthy control group).

Twenty-five (25) studies met the above-mentioned inclusion and exclusion criteria and were separated by depth cue type in Table 1. One 'mental rotation' study was also included. Although the 'mental rotation' condition is not a depth cue per se, rotating an object in depth versus in 2D-position requires mentally representing a shape in its 3D form in comparison to its 2D form. It should be noted that one study contained separate contrasts for shading-defined 3D shape and texture-defined 3D shape (Georgieva et al., 2008). Analyses were performed using either both or excluding one or the other, but no differences were found in the cluster detection. Therefore, the two contrasts were analysed concomitantly despite not being independent. A flow chart illustrating the selection and screening process can be viewed in Fig. 2. Best-practice guidelines were used to perform this meta-analysis, as recommended by Müller et al. (2018).

2.2. Activation likelihood estimation (ALE) meta-analysis method

The activation likelihood estimation (ALE) method, when combined with a thorough systematic review procedure, is a powerful tool to quantitatively assess the overlap between foci activation across studies (Eickhoff et al., 2012; Eickhoff et al., 2009, 2016). ALE is based on modeling foci as probability distributions centered at the specified coordinates that were reported in a standard space (i.e. Talairach or MNI) by individual studies. The number of subjects from each study weights the foci and the probability of activation is calculated based on both individual voxel statistical significance and a minimum cluster size threshold. With the goal of removing task-specificity from the results, only clusters of activation that are common across studies are reported. Depending on the dataset, ALE has either an equivalent or superior performance when compared to other coordinate-based meta-analytical methods (Tench et al., 2020) and image-based meta-analysis methods (Salimi-Khorshidi, Smith, Keltner, Wager and Nichols, 2009a). ALE has the advantage of having freely available software and we can include more studies since not many studies share their full statistic images, which is required for image-based meta-analysis methods (Salimi-Khorshidi, Smith, Keltner, Wager and Nichols, 2009b).

2.3. ALE based on reported coordinates of 3D versus 2D contrasts

Our coordinate-based meta-analysis of fMRI/PET studies of 3D shape

perception as compared to 2D shape perception was performed using the latest version of the GingerALE software (version 3.0.2, accessible from https://www.brainmap.org/ale/and reported by Eickhoff et al., 2009). For studies reporting their peaks in Talairach coordinates, the coordinate values were converted to MNI152 using the software's "Talairach (SPM) to MNI (Lancaster transformation)" transformation (Eickhoff et al., 2009) and saved to a target volume with a 2 mm isotropic voxel resolution. For each such volume (one per study), we then generated a dilated map of the significantly reported voxel coordinates, whereby each significant peak voxel was dilated (i.e., padded with 2 mm isotropic voxels).

Generally, ALE assesses the overlap between foci across studies by modelling them as Gaussian probability distributions centered at the peak activation coordinates (Eickhoff et al., 2012). The width of the probability distribution depends on sample size of each study, and the larger the sample the more reliable the estimation is of true activation. The maps of each study are then pooled together to identify areas of reliable foci convergence across studies. This convergence is then tested using permutation tests to ensure that the spatial convergence of foci across studies reflects an estimation of true activation and was not found by chance (Eickhoff et al., 2009, 2012). The newer version of the software also restricts the meta-analytical space to the gray matter only (Eickhoff et al., 2012) and performs random-effects analyses instead of fixed-effects analyses, by making between-experiments inferences instead of between-foci (Eickhoff et al., 2012).

For all the single dataset ALE analyses conducted in this study—full-sample and depth cue-specific analyses—cluster level threshold was set to $p < 0.001$ (family-wise error corrected at $p < 0.05$) and using 5000 permutations. For the contrast analyses, the thresholds used were 200 mm^3 as minimum cluster volume, multiple comparisons correction at $p < 0.05$ and 10000 permutations and a smoothing of 3 mm FWHM was applied.

In total, we conducted three sets of meta-analyses using GingerALE. The first was the main analysis of 3D shape perception when contrasted with 2D shape perception, regardless of depth cue. This first analysis included the full sample, with all the depth cues combined. The second set of meta-analyses consisted of five (5) sub-analyses, for the motion, disparity, texture, and shading depth cues individually. Detailed information on the number of studies, participants, and foci for each depth cue type can be retrieved from Table 2. The third subset of analysis using GingerALE aimed to investigate whether differences found across depth cues were driven by the nature of the depth cue that defined it—which was either binocular (i.e., disparity) or monocular—or by the actual shape information. Hence, we performed contrast analyses between disparity-defined 3D shapes > monocularly-defined 3D shapes (motion, texture, and shading pooled together) and the reverse monocularly-defined 3D shapes > disparity-defined 3D shapes. The results were Z-score maps of the two contrasts.

2.4. Visualization and labelling software

Significant clusters were visualized and labelled using the MNI152 template in FSLeyes (version 0.34.2). Labelling was performed using the Harvard-Oxford structural atlas available from FSL (see https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases), combined with Wang et al. (2015)'s probabilistic atlas of the visual system. Connectome Workbench 1.5.0 software (Marcus et al., 2011) was used to generate the figures, using the MSMAII.32k inflated cortical surface, along with Wang et al. (2015)'s visual cortex parcellation atlas, provided by Benson et al. (2018) (https://balsa.wustl.edu/study/show/9Zkk).

2.5. Neurosynth decoder function for reverse inference

After generation of the ALE-maps for the main full-sample meta-analysis derived from the 3D shape > 2D shape contrasts, the unthresholded ALE-map was uploaded to NeuroVault. NeuroVault is an online

Table 2
Included studies.

Authors	Year	n	Foci	Contrast	Coordinate Space	PET/ fMRI
Motion						
Schindler and Bartels	2016	15	3	motion parallax > 2D motion	MNI	fMRI
Saygin et al.,	2004	12	6	biological motion > scrambled bio motion	Talairach	fMRI
Yamamoto et al.,	2008	13	13	SKE stimulus > 2D motion	MNI	fMRI
Iwaki et al.,	2013	7	3	3D structure-from-motion (SFM) > scrambled motion	MNI	fMRI
Paradis and Droulez [N]	2000	9	4	3D SFM > scrambled motion	Talairach	fMRI
,Orban et al.,[N]	1999	20	8	3D SFM > 2D motion	MNI**	fMRI
Jastorff and Orban [N]	2009	20	6	biological motion > scrambled bio motion	MNI	fMRI
Freeman et al.,	2012	8	2	3D SFM > 2D motion	Talairach	fMRI
Peuskens et al.,	2004	15	5	3D SFM > 2D texture	MNI	fMRI
Hayashi et al., [N]	2007	10	4	motion parallax > 2D motion	MNI	fMRI
Disparity						
Faillenot et al.,	1999	7	3	3D shape-from-disparity (SFD) > 2D outlines	Talairach	PET
Nishida et al.,	2001	10	10	3D SFD > 2D outlines	MNI**	fMRI
Chen et al.,	2015	40	14	3D SFD > 2D outlines	MNI	fMRI
Ogawa et al., [N]	2013	16	16	3D movie > 2D movie	MNI	fMRI
Uji et al.,	2019	7	1	3D images > 2D images	MNI	fMRI
Gaebler et al.,	2014	26	12	3D movie > 2D movie	MNI	fMRI
Baecke et al.,	2009	26	2	3D SFD + 2D outline > 2D outlines	Talairach	fMRI
Georgieva et al., [N]	2009	22	22	3D SFD > 2D outlines	MNI	fMRI
Fortin et al.,	2002	15	5	3D SFD > 2D outlines	Talairach	PET
Durand et al., [N]	2009	20	6	3D SFD > 2D outline + 3D position	MNI	fMRI
Naganuma et al., [N]	2005	10	18	3D orientation discrimination > 2D shape discrimination	MNI	fMRI
Shading						
Georgieva et al.,*[N]	2008	18	10	3D structure-from-shading > 2D shading patterns	MNI	fMRI
Denys et al., [N]	2004	17	2	3D structure-from-shading > 2D outlines	MNI	fMRI
Texture						
Cavina-Pratesi et al., [N]	2010	8	2	3D structure-from-texture (SFT) change > 2D shape & color change	Talairach	fMRI

(continued on next page)

Table 2 (*continued*)

Authors	Year	n	Foci	Contrast	Coordinate Space	PET/ fMRI
Georgieva et al., *[N]	2008	18	10	3D SFT > 2D texture patterns	MNI	fMRI
Mental rotation						
Suchan et al.,	2006	11	3	3D shape mental rotation > 2D shape mental rotation	Talairach	fMRI

Table Note: Marked with * is a paper that reported contrasts for both shading and texture independently. Marked with ** are articles that reported Talairach, but further investigation evidenced MNI space. Marked with [N] are articles that were also included in the Neurosynth database.

open-source repository for neuroimaging studies and our results can be accessed using the link https://neurovault.org/collections/15515/. Once the data were archived in NeuroVault, we used Neurosynth's decoder function to carry out a reverse inference assessment of the cognitive, perceptual, and methodological processes most associated with the ALE map. Neurosynth (https://neurosynth.org/) is an automated meta-analytical platform that contains more than 500000 activations, reported in over 14000 neuroimaging studies that were made available through the NeuroVault repository. Neurosynth's decoder function was used to identify keywords associated with our ALE map. In total, 1307 terms with their correlation coefficients were extracted. We selected the top 50 terms by rank of correlation, after exclusion of all anatomical label terms to create a word cloud for visualization. The word cloud was created using the wordcloud function from Matlab (R2021a).

The main full-sample ALE analysis generates significant clusters of activation implicated in the 3D > 2D shape contrasts. Using Matlab's NIfTI toolbox, we treated these distinct clusters as a mask/ROI, and the ROIs were separately uploaded in NeuroVault. The Neurosynth decoder was used to estimate which ROIs were mostly associated with the terms captured in the main unthresholded ALE-map. Wordcloud terms for each cluster were thresholded at 0.1 correlation values.

3. Results

3.1. Meta-analysis on the full sample of the general 3D > 2D contrasts

An ALE meta-analysis was performed to identify reliable areas of brain activity involved in representing 3D shape information over 2D shape information regardless of depth cue. The 25 studies (Table 2) that meet inclusion and exclusion criteria during screening were analysed. The full sample comprised 25 articles, 26 contrasts, 390 participants, and 175 foci. In total, seven (7) clusters of reliable activity were detected (Table 3). Despite applying all exclusion criteria, studies were excluded only for reasons 2, 3, 4, 5, and 6 (see *literature search and study selection* section).

Those clusters included bilateral activity in diffuse higher-level visual cortex, with a more pronounced association in the right hemisphere Fig. 3a, clusters 1 and 2). Cluster 1 encompassed right hemisphere regions in the superior lateral occipital cortex, extending from the lateral occipital complex (LO1 and LO2) to V3b, V3a, TO1/hMT and the entire intraparietal sulcus (IPS0-4; Table 3). Cluster 2 extended from the inferior portion of the right lateral occipital cortex to the right inferior temporal gyrus.

Clusters 3–7 were all left hemisphere regions, including TO1/hMT, LO1, LO2, V3b and the intraparietal sulcus (IPS0-5). As in the right hemisphere, those clusters together extended from the left inferior temporal gyrus to the lateral occipital cortex and finally terminating in the parietal lobule (Fig. 3a).

Table 3 details the brain regions associated with each cluster, their average peak coordinates (MNI152 template), maximum ALE value, cluster size in volume, and the number of studies and foci that overall contributed to each cluster. In the right side of Table 3, we specified the contributions of each depth cue type for each cluster. This was computed by calculating the number of foci from each depth cue and dividing it by the total number of foci for that cluster (Table 3, converted to %; where D = disparity, M = motion, T = texture, S = shading and MR = mental rotation). From Tables 3 and it should be noted that bilateral inferior lateral occipital cortices and inferior temporal gyrus seem to receive contributions from all depth cues (except mental rotation), although the pooled studies sampled are heavily biased towards studies using disparity-defined 3D shapes.

Disparity studies showed substantial variability in stimulus type,

Table 3
Full sample ALE meta-analysis results.

Cluster #	# of peaks	Brain area(s)	Hemi	x	y	z	ALE max.	Volume (mm3)	N studies (foci)	D %	M %	T %	S %	MR %
1	9	Superior lateral occipital cortex, extending to superior parietal lobule (TO1, LO1-2, V3b, V3a, IPS0-5, SPL1)	R	33.1	−71.4	27.2	0.0308	11816	16(43)	55.8 %	27.9 %	16.3 %	–	–
2	2	Inferior lateral occipital cortex, extending to inferior temporal gyrus, temporooccipital part	R	−37.1	−76.3	9.8	0.022	3848	10(14)	42.9 %	35.7 %	7.1 %	7.1 %	–
3	3	Inferior lateral occipital cortex extending to inferior temporal gyrus, temporooccipital part	L	−32.7	−49.7	62.7	0.0188	2376	8(9)	33.3 %	44.4 %	11.1 %	11.1 %	
4	3	Superior lateral occipital cortex (IPS0-1)	L	−22.4	−61.7	58	0.0198	2048	7(9)	88.9 %	–	11.1 %	–	–
5	1	Superior parietal lobule	L	−32.7	−49.7	62.7	0.0241	1712	6(7)	71.4 %	28.6 %	–	–	–
6	2	Superior parietal lobule (IPS2)	L	−22.4	−61.8	57.9	0.025	1608	6(7)	42.9 %	42.9 %	14.3 %	–	–
7	1	Inferior lateral occipital cortex (TO1, LO1-2, V3b)	L	−36.5	−85.6	9.4	0.0143	952	3(6)	33.3 %	33.3 %	33.3 %	–	–

Table Note: Each of the seven clusters have detailed brain region labels, hemisphere (Hemi), MNI152 coordinates, ALE maxima, cluster volume in mm^3, and number of studies and foci that contributed to each cluster (IPS = intraparietal sulcus; TO1/hMT = human middle temporal visual area; V3 = third visual complex; LO = lateral occipital complex). D = disparity; M = motion; T = texture; S = shading; MR = mental rotation.

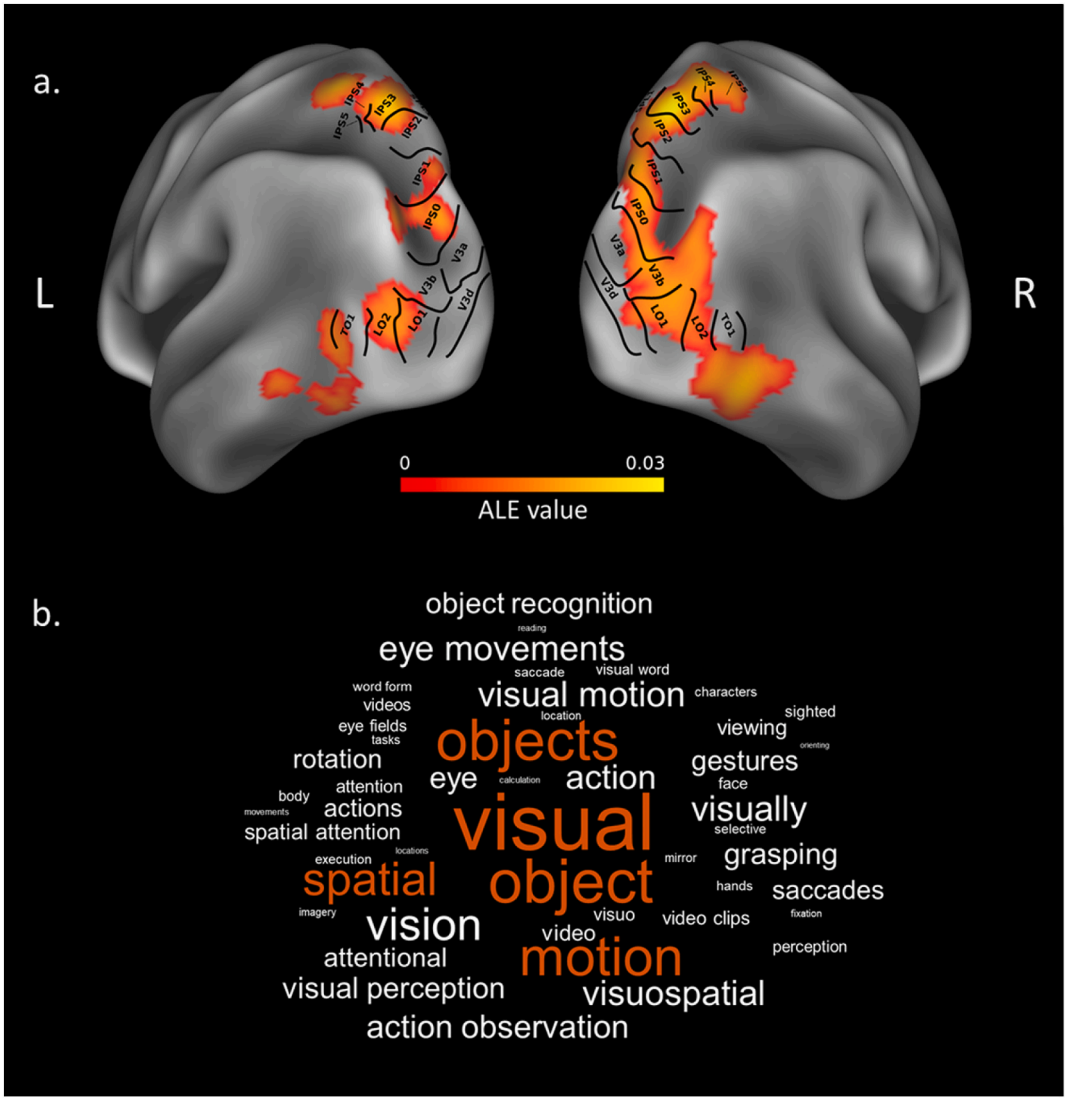


Fig. 3. ALE map reporting clusters of reliable activation in the general 3D > 2D contrasts using all studies. Clusters are numbered by size in the figure (a). Word cloud generated for reverse inference using Neurosynth decoder (b). The unthresholded ALE map was used in Neurosynth with their statistical maps and correlated with 1307 terms from their repository. Above are the first 50 conceptual terms (brain region terms removed) that correlated the most with the activation pattern of 3D shape perception over 2D shape perception. Larger fonts represent larger correlation values, the 1 % highest are represented in red. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

particularly between naturalistic 3D movies (e.g., Gaebler et al., 2014; Ogawa et al., 2013) and more controlled disparity-based stimuli. We examined the contribution of these naturalistic studies to both the full-sample and disparity-specific ALE analyses (see section below). In the full-sample analysis, Ogawa et al. (2013) contributed to clusters 1 and 3, and Gaebler et al. (2014) to cluster 2. Specifically, Ogawa contributed 2 foci out of 43 foci in cluster 1 (16 studies total), suggesting negligible influence. In cluster 3, Ogawa et al. (2013) contributed to and 1 out of 9 foci, which is more substantial. In cluster 2, Gaebler contributed 3 foci out of 14, also more substantial. However, smaller clusters (e.g., clusters 6–7) were formed with only 6–7 foci overall. Taken together, these results indicate that even if naturalistic 3D movie studies were excluded, the key clusters would remain present, though with potentially reduced ALE values.

3.1.1. Terms associated with ALE-map activations for the general 3D > 2D contrasts

The 50 highest correlation terms used to generate the word cloud, excluding brain region-related terms, can be visualized in Fig. 3b. As expected, it included conceptual terms such as ‘motion’ and ‘objects’, which relate to processing of motion depth cue and object recognition processes. The highest and lowest correlations observed were $r = 0.548$ and $r = 0.135$ for the terms “visual” and “orienting”, respectively.

In the Neurosynth decoder analysis performed on separate clusters, the cut-off was set at $r = 0.1$, with the highest r from each cluster (from 1 to 7) as being 0.210, 0.262, 0.231, 0.195, 0.244, 0.212 and 0.234, respectively (Fig. 4).

3.2. Sub-analyses

3.2.1. Meta-analysis on separate depth cues

An ALE sub-analysis was performed to identify reliable areas of brain activity involved in representing 3D shape information over 2D shape information for each individual depth cue. This was performed to assess whether the depth cues were differentially impacting the general full-sample ALE analysis. The original 25 studies were subdivided by depth cue—disparity, motion, mental rotation, texture, and shading. In total, 10 studies were found for motion, 11 for disparity, 2 for both texture and 2 shading (Table 2). One study presented independent contrasts for shading and texture (Georgieva et al., 2008), and was

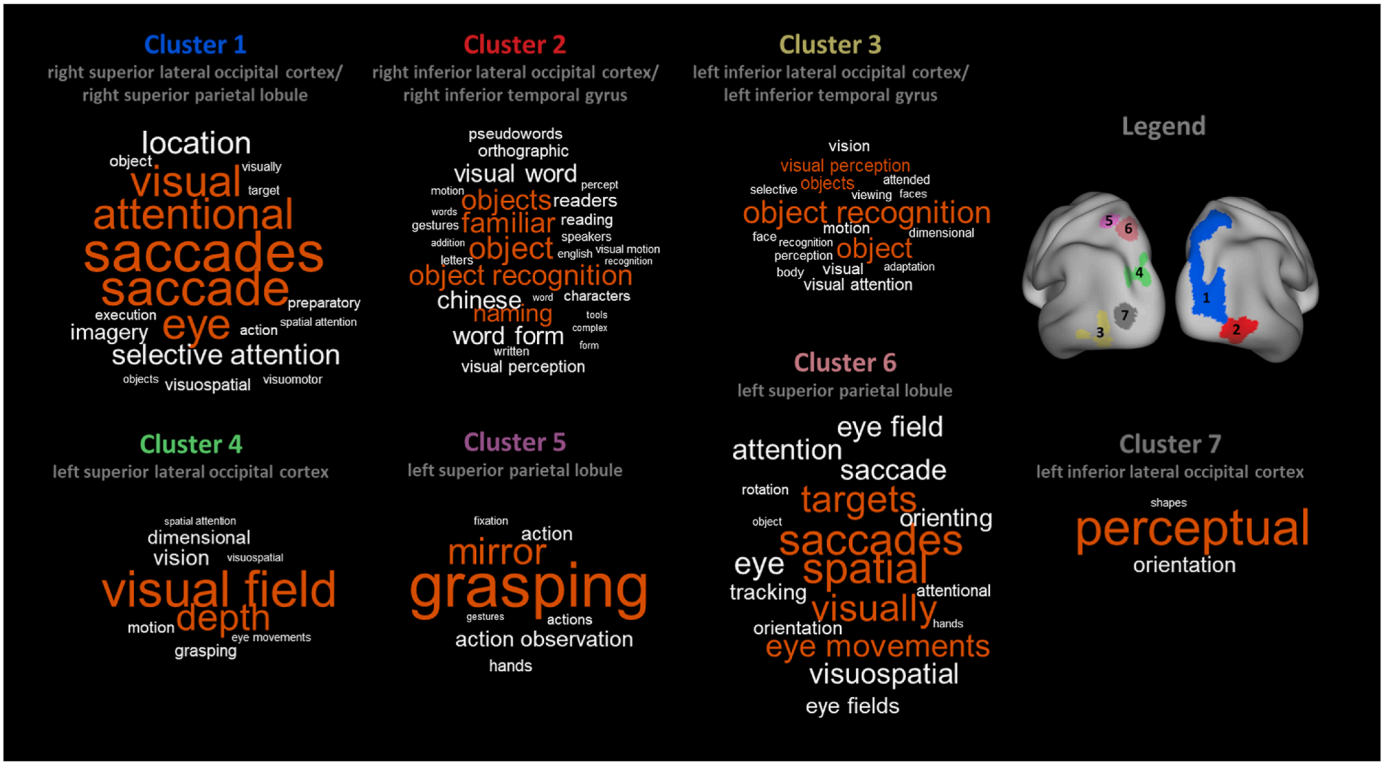


Fig. 4. Word cloud for each ALE-map cluster derived from the main 3D > 2D contrast analysis, using all studies.

reported separately with an asterisk (*).

Motion depth cue studies showed three (3) clusters of reliable activity in the right superior lateral occipital cortex, extending from areas TO1 to LO1, LO2 and V3b (Fig. 5a). Also, for motion, clusters of reliable activation were found bilaterally in the occipital fusiform gyri. The sub-analysis for disparity depth cue revealed four clusters of reliable activation. Clusters 1 and 2 were anatomically located in the right superior lateral occipital cortex, from LO1 to V3b, V3a and IPS0 (cluster 2) and extending through the intraparietal sulcus (Fig. 5b; IPS1-5, cluster 1). This activation pattern was seen to a lesser extent in the left hemisphere as well in clusters 3 and 4 (Fig. 5). Ogawa et al. (2013) was the only naturalistic study that contributed to the disparity cue-specific analysis. It contributed to 1 activation to each of the two main disparity clusters (1 out of 11 and 1 out of 12 foci, respectively). There were too few studies for texture and shading depth cues alone, and since they were underpowered, no significant clusters were found. Table 4 details the brain regions involved in motion and disparity separately, average peak coordinates, maximum ALE value, cluster volume and number of studies and foci that contributed to each cluster.

3.2.2. Meta-analytic contrast on binocularly-vs. monocularly-driven depth cues

Monocularly-defined 3D shapes (Fig. 6) reliably recruited areas in the bilateral lateral occipital cortex—ranging from hMT to lateral occipital complex and IPS for the right hemisphere (not shown but see Table 4). To test whether the differences in areas of convergence observed across the depth cues were driven by binocularly-defined 3D shapes instead of 3D shape perception—i.e., lower-level stereoscopic feature processing versus shape content or object information processing—we performed a sub-analysis contrasting binocularly-defined 3D shapes with monocularly-defined 3D shapes. Results show that dorsal stream areas, ranging from bilateral IPS0-1 and superior parietal lobules and right IPS2-5 were significantly more implicated by binocularly-defined 3D shapes when contrasted to monocularly-defined 3D shapes (Fig. 6a). In addition, the monocularly-defined 3D shapes condition showed significant spatial convergence in the left lateral occipital cortex (Fig. 6b). The conjunction analyses (binocularly-defined 3D shapes ∩ monocularly-defined 3D shapes) showed spatial convergence in the anterior portion of the right superior lateral occipital cortex, from V3b to IPS0-4 Fig. 6c).

4. Discussion

The main ALE meta-analysis of neuroimaging studies shows that 3D-shape information is represented throughout the high-level visual cortex, in both ventral and dorsal pathways when all depth cues are jointly considered. Although the sub-analyses were underpowered, we found trends suggesting depth-cue specificities in 3D-shape processing—particularly that of motion-defined 3D-shapes are more dominantly represented in the right dorsal and ventral streams, whereas disparity-defined 3D-shapes predominantly recruit higher-level areas in the bilateral dorsal stream. The second set of sub-analyses that contrasted binocularly-driven with monocularly-driven 3D shapes showed that bilateral dorsal stream is more strongly recruited for disparity-defined 3D shapes whereas left ventral stream is more strongly recruited by monocularly-defined 3D shapes. Lastly, based on the conjunction analysis, we found that the right ventral and dorsal streams together represent monocular and binocularly-driven 3D shape information. Together, these findings imply that (1) integration between the two streams is likely to be subserving 3D shape processing; (2) there may be segregation in the cortical substrates of 3D shape processing depending on the depth cue used to represent the objects; and (3) 3D shape processing may be lateralized depending on binocularity.

There is accumulating evidence that object-related information is also encoded in the dorsal stream and that the two streams may be integrated especially in the context of object perception from 3D depth cues. The two streams possess extensive anatomical connections (Takemura et al., 2015) and the dorsal stream jointly represents structural information of complex objects (Freud et al., 2017; Freud et al., 2018; Jeong and Xu, 2016). Although in the present meta-analysis we

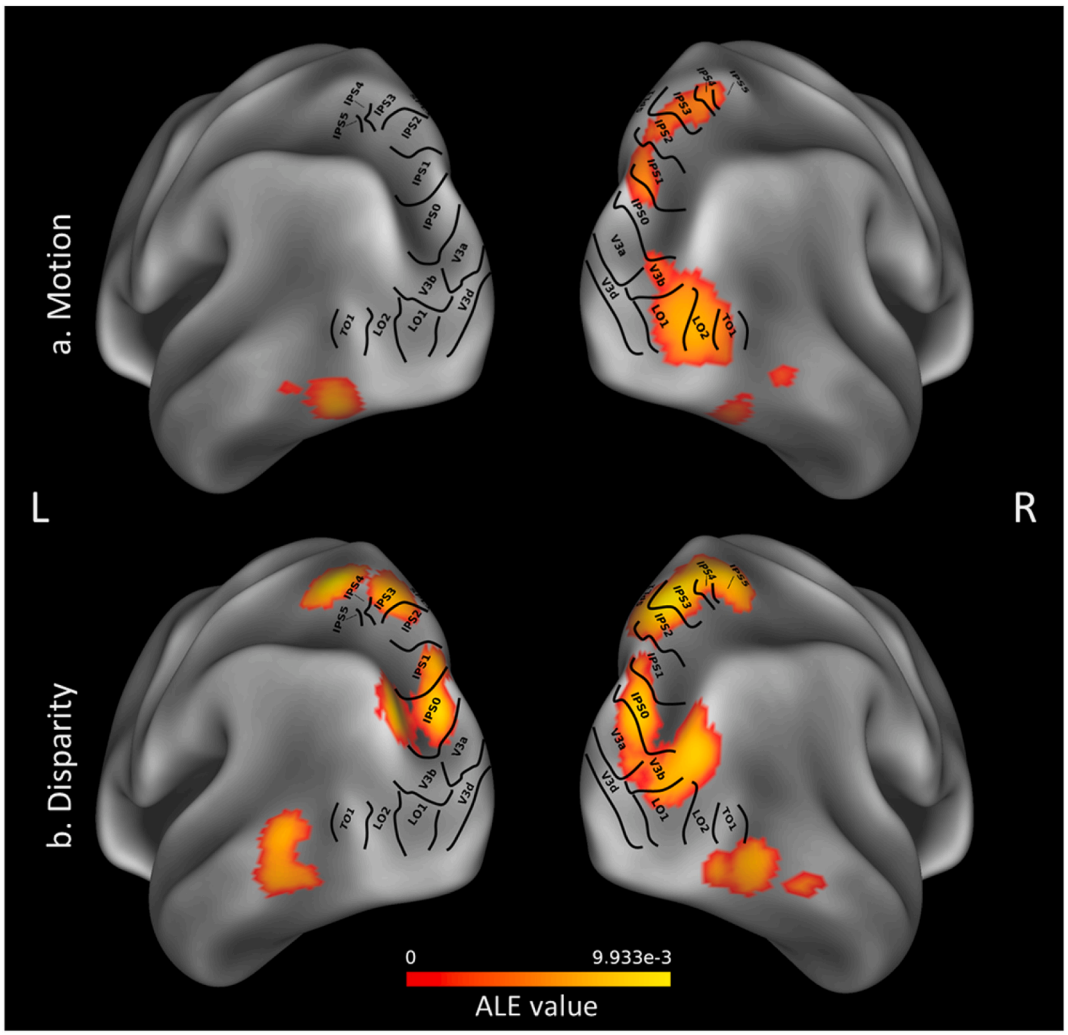


Fig. 5. ALE map reporting clusters of reliable activation for motion-defined 3D shapes (A) and disparity-defined 3D shapes (B). Clusters are numbered by size in the figure (IPS = intraparietal sulcus; TO1/hMT = human middle temporal visual area; V3 = third visual complex; LO = lateral occipital complex). More details in Table 4.

did not directly address integration or functional connectivity across visual areas underlying 3D shape perception, the reliable widespread activation networks observed indicate that the two streams are jointly involved in processing 3D information of shapes. The combined efforts of these studies, together with the first major claim from our meta-analytic results, suggest that dorsal and ventral integration is needed for 3D object recognition (for reviews, see Farivar, 2009; Freud et al., 2016, 2020).

ALE values for the main analysis (Fig. 3) are higher in dorsal stream regions (IPS) when compared to the ventral stream regions (LOC). This reiterates the dorsal stream's role in processing 3D shapes defined by depth cues when contrasted to their 2D counterparts. In fact, it could implicate other known functionalities of the dorsal stream, such as processing 3D structure to enable object grasping and manipulation (Janssen and Scherberger, 2015) and even the fact that 3D shapes contain more attention-grabbing information, which is a facet of attention-related functionalities of the parietal cortex (for a review, see Behrmann et al., 2004). However, because of the heterogeneity of the stimuli comprised in the current analyses, we can argue that the 'surviving' mechanism across studies is the 3D shape processing aspect, considering it as the commonality between them. This is reiterated by recent evidence showing dissociation between attentional and shape processing mechanisms in the dorsal stream (Goldstein-Marcusohn et al., 2024). Goldstein-Marcusohn et al. (2024) remarked that attentional demands modulated shape sensitivity similarly in both ventral and dorsal streams, despite attentional processes occurring in overlapping dorsal regions.

A previous ALE meta-analysis addressed a related question (Farshchi, 2022). However, their main objective was not specifically to investigate depth-cue representation of 3D shape vs 2D shape. For this reason, their search strategy did not include terms related to depth cues. Farshchi (2022) used few search terms and also fewer databases which may have contributed to undersampling of relevant studies (i.e., 199 records after deduping; 16 articles included in the meta-analysis). Consequently, it is difficult to determine the extent to which their results reflect depth cue-tolerant 3D shape activations. Another difference was their focus on task complexity in secondary analyses (e.g., passive versus judgment tasks). They also adopted a different conceptual approach to contrast selection (e.g., including scrambled images with no visible 2D shape or contour as a control). We speculate that these divergences likely contributed to the discrepancies in study inclusion between the two meta-analyses, although several studies overlap. Overall, their main ALE analysis results (full sample) are comparable to ours, showing consistent recruitment of high-level ventral and dorsal visual areas, but with some differences, considering we additionally observed engagement of the left superior parietal lobule, which was not reported in their study.

Although our depth cue-specific sub-analyses for motion and disparity were underpowered, the trend observed here for depth cue-

Table 4
Results of the ALE sub-analysis.

Cluster #	# of peaks	Brain area(s)	Hemi	x	y	z	ALE max.	Volume (mm3)	N studies (foci)
		Motion							
1	2	Superior lateral occipital cortex (TO1, LO1-2, V3b)	R	41.4	−80.8	8.1	0.0064	3152	4(7)
2	2	Inferior lateral occipital cortex extending to occipital fusiform gyrus	R	47	−66.3	−15.4	0.0064	2136	4(4)
3	3	Superior lateral occipital cortex extending to superior parietal lobule (IPS0-5)	R	25.5	−69.4	48.9	0.0048	1960	5(6)
4	1	Inferior lateral occipital cortex extending to occipital fusiform gyrus	L	−46.1	−68.9	11.5	0.0081	1896	3(3)
		Disparity							
1	6	Superior lateral occipital cortex extending to superior parietal lobule (IPS1-5, SPL1)	R	28.8	−68.8	45	0.0107	12960	10(27)
2	2	Superior lateral occipital cortex (V3a, IPS0-2)	L	−25.3	−79.8	35.6	0.0108	4128	6(8)
3	2	Superior lateral occipital cortex extending to superior parietal lobule (IPS2-4, SPL1)	L	−28.7	−53.5	61	0.0101	3904	5(8)
4	5	Inferior lateral occipital cortex (TO1)	R	48.4	−67.6	−6	0.0066	2064	4(6)
5	2	Inferior lateral occipital cortex	L	−52.3	−64	−0.1	0.0075	2064	5(5)
		Monocularly-defined 3D shapes							
1	5	Lateral occipital cortex (V3d, LO1, V3b, V3a)	R	41.9	−75.3	−0.4	0.0094	8984	6(17)
2	3	Lateral occipital cortex (LO1-2, V3b)	L	−42.2	−79.2	−1.7	0.0105	7136	6(14)
3	4	Superior lateral occipital cortex extending to superior parietal lobule (IPS0-4)	R	24.3	−67.5	50.6	0.0073	4688	6(10)
		Disparity- ∩ Monocularly-defined 3D shapes							
1	2	Superior lateral occipital cortex extending to superior parietal lobule (IPS2-4)	R	24.6	−60.5	7.7	0.0072	1992	8(9)
2	3	Inferior lateral occipital cortex	R	48.2	−68	−6.7	0.0063	1504	7(7)
3	2	Superior lateral occipital cortex (LO1, V3b, V3a, IPS0)	R	34.5	−83	16.5	0.0058	896	2(2)
4	2	Superior lateral occipital cortex (IPS0-1)	R	26.6	−77.3	39.8	0.0054	600	4(4)
		Disparity- > Monocularly-defined 3D shapes							
1	9	Superior lateral occipital cortex (IPS0-1)	L	−25.2	−80.1	37.2	0.0009	3144	6(7)
2	3	Superior lateral occipital cortex (V3b, V3a, IPS0)	R	30.5	−84.6	27.2	0.0074	1368	2(2)
3	1	Inferior lateral occipital cortex	L	−53	−63.6	1.1	0.0186	720	2(2)
4	4	Superior parietal lobule	R	37.5	−48.6	61.9	0.0134	704	1(1)
5	3	Superior lateral occipital cortex IPS3	R	24.5	−65.2	62.4	0.0136	632	1(1)
6	2	Superior lateral occipital cortex IPS2	R	13.9	−64.6	48.1	0.0346	328	1(1)
		Monocularly- > Disparity-defined 3D shapes							
1	5	Inferior lateral occipital cortex extending to occipital fusiform gyrus	L	−47.7	−74.8	−11.3	0.026	1384	2(2)
2	4	Inferior lateral occipital cortex (TO1, LO2)	L	−42.5	−90	−5.6	0.0085	496	1(1)

Table Notes: Each cluster has detailed brain region labels, hemisphere (Hemi), MNI152 coordinates, ALE maxima, volume in mm^3, and number of studies and foci that contributed to each cluster (IPS = intraparietal sulcus; hMT = human middle temporal visual area; V3 = third visual complex; LO = lateral occipital complex).

dependent cortical segregation for 3D vs. 2D corroborates a recent patient study on posterior cortical atrophy patients. These patients experience deficits in 3D perception that relates to atrophy and reduced fMRI response in ventral and dorsal pathways (Gillebert et al., 2015) depending on depth-cue. Namely, 3D shape processing deficits for shaded objects were associated with volume loss in the ventral stream while dorsal stream volume loss was associated with deficits for disparity-defined 3D shapes. It is important to note that the disparity studies included in the meta-analysis varied in stimulus type, incorporating both naturalistic 3D movies and more controlled disparity stimuli. The overall contribution of naturalistic studies to cluster formation was relatively small. Therefore, excluding these studies would not alter the key dorsal activations (e.g., in IPS), but might only reduce ALE values slightly. This variability further motivated our secondary contrast analyses (binocularly-versus monocularly-defined 3D shapes), which aimed to disentangle disparity-specific processing from shape-from-disparity representations.

Bilateral IPS was implicated in the contrast of disparity-defined 3D shapes with monocularly-defined 3D shapes, meaning that the bilateral IPS association observed in the main analysis is at least partially driven by disparity processing. This is in accordance with the binocular disparity processing network, which is widespread throughout ventral and dorsal visual streams, including strong activations in the IPS dorsally (Abed Rabbo et al., 2018; Tsao et al., 2003; Verhoef et al., 2016). However, it is not possible to completely disregard the role of the IPS for 3D shape processing, especially because the conjunction analyses (disparity-defined 3D shapes ∩ monocularly-defined 3D shapes condition) showed spatial convergence in the right IPS. This is in accordance with evidence from ROI-based studies showing complex and behaviourally relevant object representations in anterior IPS (Freud et al., 2017; Jeong and Xu, 2016). It could also imply lateralization of 3D shape processing when compared to binocularly-defined 3D shapes. More specifically, the left-hemisphere being more sensitive to monocularly-defined 3D shapes, whereas the right hemisphere may be more dedicated to the processing of 3D shape derived from disparity. Unfortunately, because of the limited number of studies for most monocularly-defined depth cues (i.e., texture and shading), we were unable to test depth-cue-specific activations for those depth cues.

The Neurosynth analysis reinforced the view that the regions recruited in the main analysis are related to visual object recognition concepts. In the individual cluster analysis, superior parietal lobule clusters were predominantly associated with oculomotor control, attentional processes, and action behavior concepts. This is in accordance with the functional capacities of the superior parietal lobule (Behrmann et al., 2004; Fogassi and Luppino, 2005; Gamberini et al., 2021). Concepts relating to object perception were more associated with inferior temporal gyrus and inferior lateral occipital cortex, ventral areas traditionally associated with visual object recognition processes

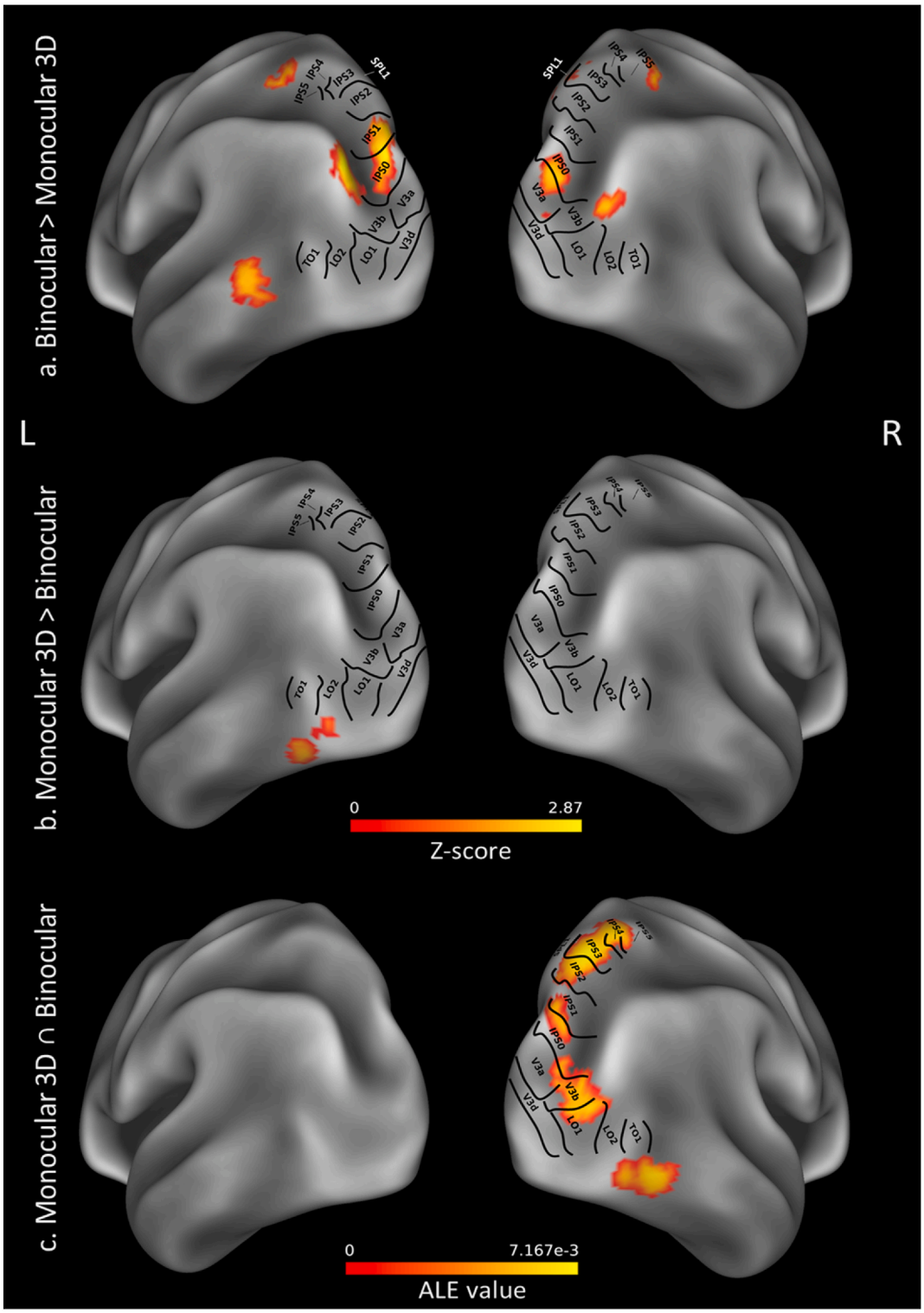


Fig. 6. Z-score maps reporting clusters of reliable activation for monocularly-defined 3D shapes > binocularly-defined 3D shapes (a.) binocularly-defined 3D shapes > monocularly-defined 3D shapes (b.) contrasts. ALE map for the conjunction of activation between binocularly- and monocularly-defined 3D shapes conditions (c.). The Wang et al. (2015) atlas was used for visualization. (IPS = intraparietal sulcus; TO1/hMT = human middle temporal visual area; V3 = third visual complex; LO = lateral occipital complex). More details in Table 4.

(Bracci et al., 2017; Grill-Spector and Weiner, 2014; Peissig and Tarr, 2006). However, our results show that activity of the ventral stream alone does not paint a complete picture when considering 3D visual object recognition. These concepts are likely to have risen due to the traditional use of 2D-feature based objects as a proxy for visual object recognition in many studies (Andresen et al., 2009; Bernstein et al., 2014; Bracci and Op de Beeck, 2016; Bracci et al., 2019; Konkle and Oliva, 2012; Mocz et al., 2023). It is worth noting that 11 out of the 25 studies (corresponding to 12 of the 26 independent contrasts) included in our meta-analysis are also part of the Neurosynth database (Table 2). Their distribution was balanced across the distinct depth cues (4 for motion, 4 for disparity, 2 for texture, and 2 for shading). Although this

potentially introduces circularity to the analysis, the impact is expected to be minimal given that the Neurosynth database contains over 14 thousand fMRI studies. In addition, because the overlap is small and evenly distributed across depth cues, any circularity would affect all cue categories similarly, thus not biasing comparisons between them and their individual contributions. The reverse-inference maps in Neurosynth are derived from large-scale, topic-based meta-analytic coactivation patterns rather than direct inclusion of individual foci. Therefore, their influence from a small subset of studies is expected to be negligible.

Several theoretical and computational models of visual object recognition have been proposed over the years. They largely differ on various aspects, including whether the extracted features from the viewed images are primarily 2D or 3D and how they approach the issue of object variability across views—by using stored 2D or 3D templates of objects. Historically, most object recognition computational models are based on the architecture and functionality of the ventral visual stream alone. They generally rely on 2D feature and 2D pattern extractions from the viewed image—e.g., edge extractions—and generation of 2D-based intermediate representations derived from learned correlations between derived hierarchical features, which are then used as templates for visual object recognition (Poggio and Edelman, 1990; Poggio and Riesenhuber, 1999; Ullman and Basri, 1991).

The 2D feature-based family of models have been supported by behavioral and neurophysiological research at the time (Bülthoff and Edelman, 1992; Edelman and Bülthoff, 1992; Jolicoeur, 1985; Logothetis and Pauls, 1995; Tarr and Pinker, 1989). Edelman and Bülthoff (1992) showed that when humans were trained in a specific view of a novel wireframe object, their error rates in object identification performance increased with rotation angle in depth from the learned view. Those results were replicated for nonhuman primates (Logothetis et al., 1994), but when animals were trained with as few as three different views of the same object, they could effectively recognize the object in untrained views (Logothetis and Pauls, 1995). Their results supported the use of interpolation of stored 2D object templates to generalize for novel views for visual object recognition (Poggio and Edelman, 1990; Ullman and Basri, 1991). Later, neurophysiological studies have shown that a single IT neuron in the macaque can have object selectivity, regardless of its size or position on the retinal image (Grill-Spector et al., 2001; Grill-Spector et al., 1998; Riesenhuber and Poggio, 1997) but that they were also tuned to viewpoint (Logothetis and Pauls, 1995), again supporting 2D based-models of object recognition.

However, ventral areas that have been traditionally associated with the processing of 2D shape information are also reliably representing 3D shape information. This is in accordance with the previously mentioned electrophysiological studies in nonhuman primates which showed that 3D features of objects are being explicitly encoded throughout the ventral cortex hierarchy (Hinkle and Connor, 2002; Hung et al., 2012; Srinath et al., 2021; Vaziri et al., 2014; Yamane et al., 2008) and further supported by our ALE meta-analytic results in humans. These studies together demonstrate that object representations in ventral stream are not only 2D-feature based, but also 3D-feature based.

2D feature-based computational models are mainly modelled after the ventral stream (Cadieu et al., 2014; Christensen and Zylberberg, 2020; Kar et al., 2019; Poggio and Riesenhuber, 1999; Riesenhuber and Poggio, 2000; Yamins et al., 2014). In addition, with current advances of deep-convolutional neural networks (DNNs), recent studies directly test correspondences between the hierarchical layers of the ventral pathway and the hierarchical layers of DNNs (Güçlü and van Gerven, 2015)—i.e., general purpose computer vision models—as a proxy of human visual object recognition, while other DNNs were explicitly created for aiming for biological plausibility and are directly inspired by the primate ventral pathway (Kubilius et al., 2018). Although modern DNN models can reach human object recognition performance (Mur et al., 2013; Rajalingham et al., 2015), they fail to generalize when humans do, especially when images are particularly tweaked. Su and colleagues (2019) have shown that by changing a single pixel in a test image, performance in three distinct DNNs falls considerably—misclassifying in average 73.22 % of the time—although the objects remain effortlessly recognizable by humans. Similar results are observed when object images are purposefully distorted with blurring or noise (Dodge and Karam, 2017; Geirhos et al., 2018). If 3-Dimensionality of objects, such as overall object surface, medial axis and structural relationships between object parts were taken into consideration in such models, they would not have bluntly failed—e.g., medial axis and relationship between parts and articulations do not change with additive noise or blur. Most strikingly, considering the role of the dorsal stream in 3D object processing (Ayzenberg and Behrmann, 2022; Freud et al., 2017; Jeong and Xu, 2016) and depth cue processing and integration (Dövencioğlu et al., 2013; Murphy et al., 2013) mentioned earlier and considering the extensive involvement of the dorsal stream in 3D shape processing, observed in the current meta-analysis, visual object recognition models should account for the dorsal stream contribution.

There is increasing evidence that object perception relies on a surface-based code that integrates 2D shape-based information with 3D structural and configural information that is used jointly for object recognition (Arguin et al., 2019; Kim et al., 2019; Leek et al., 2005; Pasupathy et al., 2019; Yamane et al., 2008). However, even though 3-Dimensionality is important for object recognition, it is difficult to probe 3-Dness without 2-Dness—or distinguishing between 2D and 3D representations. In Leek et al.'s, (2005) study which suggests a surface-based model of 3D shape representations, they used 2D line-drawing contours to delineate surfaces and volumetric components, therefore not eliminating the 2D cue contributions. Their "open contour" stimuli, also 2D-based but composed with noncontiguous edges of the original shape—or an "interrupted surface"—could for example be harder to perceive simply because of good continuation—i.e., identification is harder when line segments present poorer continuation (Pelli et al., 2009). Therefore, because 2D representations can be derived from 3D projections, they are difficult to separate. It is possible that 3D shapes simply contain more object information that can be relevant for perception.

By using same-identity shapes rendered by unique depth cues it is also possible to probe 3Dness over 2Dness—namely, by studying depth cue invariance. The nature of the lower-level 2D features can be extraordinarily different although their identity (Dehmoobadsharifabadi and Farivar, 2016) or category (Akhavein, 2017; Akhavein et al., 2018) are still perceived as being the same. Some depth cues are 3D by nature, such as the case for objects rendered using binocular disparity or structure-from-motion which are edgeless. Depth cue invariance suggests that a common dimension is extracted from those objects regardless of depth cue, and this dimension is essential for recognition, such as their surface or other influential 3D structural information (e.g., relative distances between meaningful object parts, or medial axis extraction) as information. Therefore, future endeavours in understanding how 2Dness and 3Dness intertwine in object recognition will benefit from studies of depth cue invariance.

4.1. Limitations

The ALE method has some limitations. They include not being able to incorporate studies without whole-brain acquisition—which can be common in hypotheses-driven vision studies, considering a higher voxel resolution can be acquired in the visual cortex when using smaller fields-of-view. We also were not able to include studies that performed analysis on the cortical surface—i.e. not the volume—which is increasingly more common in recent studies, considering alignment across subjects is more reliable in surface-based studies when compared to volume-based ones (Coalson et al., 2018; Fischl et al., 1999). We are also unable to include studies that perform multivoxel pattern analysis (MPVA) as ALE requires peak coordinate activations reported from the more traditional, whole-brain univariate analyses.

ALE can pinpoint large-scale networks that are more dominantly activated in specific contrasts (higher BOLD signal) and can therefore reveal large-scale networks that are overall involved in a subjacent cognitive process. With MVPA, it's possible for example to assess subtle differences between 3D and 2D processing, or that a same area is responding to both 2D and 3D shape information but with distinct activation patterns, even though signal amplitude for 3D is larger (i.e., 3D containing more information, being more salient or interesting, etc.). In the perspective of large-scale, whole-brain organization however, ALE is an appropriate tool to evidentiate areas that are consistently recruited by depth cue defined 3D shapes, despite task or stimuli specificity (or heterogeneity in design). This is even more pressing when considering that univariate analyses—and reporting peak coordinates—have been performed since decades of neuroimaging research. Only in recent years, with the Open Science movement, researchers have begun to openly share their raw experimental data and full statistic images that could be then used for more refined and specialized analysis and this should be pursued in future studies to assess the intricate contributions of distinct depth cues for 3D object recognition.

It is also worth noting that the sub-analyses (i.e., disparity, motion and the contrast analyses) are underpowered, with only 11 studies for disparity (<16 studies are not recommended, see Eickhoff et al., 2012), thus inference here has to be done with caution.

5. Conclusion

Overall, the results of this meta-analysis support representations of 3D shapes throughout the high-level visual cortices, in both ventral and dorsal pathways and regardless of the depth cue that defines them. The trends indicate that binocularly-defined 3D shapes may be more prominently represented in the bilateral dorsal stream, whereas monocularly-defined 3D shapes are more dominantly represented in the left ventral stream. These observations point to a possible degree of segregation and depth cue specificity in 3D shape processing, as suggested by motion- and disparity-defined 3D shapes. These findings are consistent with the notion that depth cue integration between the two streams likely subserves complex shape processing, specifically by representing 3D structure of objects and endorsing 3D shape perception.

CRediT authorship contribution statement

Luiza P. Volpi: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **R. Nathan Spreng:** Writing – review & editing, Supervision, Methodology. **Reza Farivar:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank the Canadian Institutes of Health Research (CIHR) funds (#29517) that supported this study.

Data availability

The ALE maps data can be accessed on https://identifiers.org/neurovault.collection:15515. For specific codes (i.e., word cloud main and cluster analyses) and raw coordinate points data, please contact the authors.

References

Abed Rabbo, F., Koch, G., Lefèvre, C., Seizeur, R., 2018. Stereoscopic visual area connectivity: a diffusion tensor imaging study. Surg. Radiol. Anat. 40 (11), 1197–1208. https://doi.org/10.1007/s00276-018-2076-3.

Akhavein, H., 2017. Depth-Cue Invariant Object Representations in the Visual Cortex. (April).

Akhavein, H., Dehmoobadsharifabadi, A., Farivar, R., 2018. Magnetoencephalography adaptation reveals depth-cue-invariant object representations in the visual cortex. J. Vis. 18 (12), 6. https://doi.org/10.1167/18.12.6.

Andresen, D.R., Vinberg, J., Grill-Spector, K., 2009. The representation of object viewpoint in human visual cortex. Neuroimage 45 (2), 522–536. https://doi.org/10.1016/j.neuroimage.2008.11.009.

Arguin, M., Marleau, I., Aubin, M., Zahabi, S., Leek, E.C., 2019. A surface-based code contributes to visual shape perception. J. Vis. 19 (11), 1–23. https://doi.org/10.1167/19.11.6.

Ayzenberg, V., Behrmann, M., 2022. The dorsal visual pathway represents object-centered spatial relations for object recognition. J. Neurosci. 42 (23), 4693–4710. https://doi.org/10.1523/JNEUROSCI.2257-21.2022.

Baecke, S., Lützkendorf, R., Tempelmann, C., Müller, C., Adolf, D., Scholz, M., Bernarding, J., 2009. Event-Related Functional Magnetic Resonance Imaging (Efmri) of depth-by-disparity Perception : Additional Evidence for right-hemispheric Lateralization, vols. 453–458. https://doi.org/10.1007/s00221-009-1844-z.

Behrmann, M., Geng, J.J., Shomstein, S., 2004. Parietal cortex and attention. Curr. Opin. Neurobiol. 14 (2), 212–217. https://doi.org/10.1016/j.conb.2004.03.012.

Benson, N.C., Jamison, K.W., Arcaro, M.J., Vu, A.T., Matthew, F., Timothy, S., et al., 2018. The Human Connectome Project 7 Tesla Retinotopy Dataset : Description and Population Receptive Field Analysis Essen, vol. 18, pp. 1–22.

Bernstein, M., Oron, J., Sadeh, B., Yovel, G., 2014. An integrated face–body representation in the fusiform gyrus but not the lateral Occipital Cortex Michal. J. Cognit. Neurosci. 26 (11), 2469–2478. https://doi.org/10.1162/jocn.

Bracci, S., Op de Beeck, H., 2016. Dissociations and associations between shape and category representations in the two visual pathways. J. Neurosci. 36 (2), 432–444. https://doi.org/10.1523/JNEUROSCI.2314-15.2016.

Bracci, S., Ritchie, J.B., de Beeck, H.O., 2017. On the partnership between neural representations of object categories and visual features in the ventral visual pathway. Neuropsychologia 105 (June), 153–164. https://doi.org/10.1016/j.neuropsychologia.2017.06.010.

Bracci, S., Ritchie, J.B., Kalfas, I., Op de Beeck, H.P., 2019. The ventral visual pathway represents over animacy, unlike human behavior and deep neural networks. J. Neurosci. 39 (33), 6513–6525.

Bramer, W.M., Giustini, D., Jonge, G. B. De, Holland, L., Bekhuis, T., 2016. De-duplication of database search results for systematic reviews in EndNote, 104 (July), 240–243.

Bülthoff, H.H., Edelman, S., 1992. Psychophysical support for a two-dimensional view interpolation theory of object recognition. Proc. Natl. Acad. Sci. U. S. A. 89 (1), 60–64. https://doi.org/10.1073/pnas.89.1.60.

Cadieu, C.F., Hong, H., Yamins, D.L.K., Pinto, N., Ardila, D., Solomon, E.A., et al., 2014. Deep neural networks rival the representation of primate IT cortex for core visual object recognition. PLoS Comput. Biol. 10 (12). https://doi.org/10.1371/journal.pcbi.1003963.

Cavina-Pratesi, C., Kentridge, R.W., Heywood, C.A., Milner, A.D., 2010. Separate channels for processing form , texture , and color : evidence from fMRI adaptation and visual object agnosia. Cerebr. Cortex (October), 2319–2332. https://doi.org/10.1093/cercor/bhp298.

Chainay, H., Humphreys, G.W., Humphreys, G.W., 2010. The real-object advantage in agnosia : evidence for a role of surface and depth information in object recognition T HE REAL - OBJECT ADVANTAGE IN AGNOSIA : 3294. https://doi.org/10.1080/02643290042000062.

Chen, C., Wang, J., Li, K., Liu, Y., Chen, X., 2015. Visual fatigue caused by watching 3DTV : an fMRI study, 14 (Suppl. 1), 1–12.

Christensen, E., Zylberberg, J., 2020. Models of the ventral stream that categorize and visualize images. bioRxiv. https://doi.org/10.1101/2020.02.21.958488, 2020.02.21.958488.

Coalson, T.S., Van Essen, D.C., Glasser, M.F., 2018. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. Proc. Natl. Acad. Sci. U. S. A. 115 (27), E6356–E6365. https://doi.org/10.1073/pnas.1801582115.

Dehmoobadsharifabadi, A., Farivar, R., 2016. Are face representations depth cue invariant? J. Vis. 16 (8), 6. https://doi.org/10.1167/16.8.6.

Denys, K., Vanduffel, W., Fize, D., Nelissen, K., Peuskens, H., Essen, D. Van, Orban, G.A., 2004. The Processing of Visual Shape in the Cerebral Cortex of Human and Nonhuman Primates : A Functional Magnetic Resonance Imaging Study. 24 (10), 2551–2565. https://doi.org/10.1523/JNEUROSCI.3569-03.2004.

Dodge, S., Karam, L., 2017. A Study and Comparison of Human and Deep Learning Recognition Performance Under Visual Distortions.

Dövencioğlu, D., Ban, H., Schofield, A.J., Welchman, A.E., 2013. Perceptual integration for qualitatively different 3D cues in the human brain. J. Cognit. Neurosci. 25 (9), 1527–1541. https://doi.org/10.1162/jocn.

Durand, J.B., Peeters, R., Norman, J.F., Todd, J.T., Orban, G.A., 2009. Parietal regions processing visual 3D shape extracted from disparity. Neuroimage 46 (4), 1114–1126. https://doi.org/10.1016/j.neuroimage.2009.03.023.

Edelman, S., Bülthoff, H.H., 1992. Orientation dependence in the recognition of familiar and novel views of threedimensional objects. Vis. Res. 32 (12), 2385–2400. https://doi.org/10.1016/0042-6989(92)90102-O.

Eickhoff, S.B., Bzdok, D., Laird, A.R., Kurth, F., Fox, P.T., 2012. NeuroImage activation likelihood estimation meta-analysis revisited. Neuroimage 59 (3), 2349–2361. https://doi.org/10.1016/j.neuroimage.2011.09.017.
Eickhoff, S.B., Laird, A.R., Grefkes, C., Wang, L.E., Zilles, K., Fox, P.T., 2009. Coordinate-Based Activation Likelihood Estimation Meta-Analysis of Neuroimaging Data : a Random-Effects Approach Based on Empirical Estimates of Spatial Uncertainty, pp. 2907–2926. https://doi.org/10.1002/hbm.20718, 2926(November 2008).
Eickhoff, S.B., Nichols, T.E., Laird, A.R., Hoffstaedter, F., Amunts, K., Fox, P.T., et al., 2016. Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. Neuroimage 137, 70–85. https://doi.org/10.1016/j.neuroimage.2016.04.072.
Faillenot, I., Decety, J., Jeannerod, M., 1999. Human Brain Activity Related to the Perception of Spatial Features of Objects, 124, pp. 114–124.
Farivar, R., 2009. Dorsal-ventral integration in object recognition. Brain Res. Rev. 61 (2), 144–153. https://doi.org/10.1016/j.brainresrev.2009.05.006.
Farivar, R., Blanke, O., Chaudhuri, A., 2009. Dorsal-ventral integration in the recognition of motion-defined unfamiliar faces. J. Neurosci. 29 (16), 5336–5342. https://doi.org/10.1523/JNEUROSCI.4978-08.2009.
Farshchi, M., 2022. Neural substrates that maintain perceiving 3D information: an ale meta-analysis study. Psychol-J High Sch E 19 (4), 663–683. https://doi.org/10.17323/1813-8918-2022-4-663-683.
Fischl, B., Sereno, M.I., Tootell, R.B.H., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp. 8, 272–284.
Fogassi, L., Luppino, G., 2005. Motor functions of the parietal lobe. Curr. Opin. Neurobiol. 15 (6), 626–631. https://doi.org/10.1016/j.conb.2005.10.015.
Fortin, A., Ptito, A., Faubert, J., Ptito, M., 2002. Cortical areas mediating stereopsis in the human brain : a PET study, 13 (6), 895–898.
Freeman, E.D., Sterzer, P., Driver, J., 2012. fMRI correlates of subjective reversals in ambiguous structure-from-motion. J. Vis. 12 (6), 1–19. https://doi.org/10.1167/12.6.35.Introduction.
Freud, E., Behrmann, M., Snow, J.C., 2020. What does dorsal cortex contribute to perception? Open Mind 4, 40–56. https://doi.org/10.1162/opmi_a_00033.
Freud, E., Ganel, T., Shelef, I., Hammer, M.D., Avidan, G., Behrmann, M., 2017. Three-dimensional representations of objects in Dorsal Cortex are dissociable from those in ventral Cortex. Cerebr. Cortex 27 (1), 422–434. https://doi.org/10.1093/cercor/bhv229.
Freud, E., Plaut, D.C., Behrmann, M., 2016. "What" is happening in the dorsal visual pathway. Trends Cognit. Sci. 20 (10), 773–784. https://doi.org/10.1016/j.tics.2016.08.003.
Freud, E., Robinson, A.K., Behrmann, M., 2018. More than action: the dorsal pathway contributes to the perception of 3-D structure. J. Cognit. Neurosci. 30 (7), 1047–1058. https://doi.org/10.1162/jocn.
Gaebler, M., Biessmann, F., Lamke, J., Müller, K., Walter, H., Hetzer, S., 2014. Stereoscopic depth increases intersubject correlations of brain networks. Neuroimage 100, 427–434. https://doi.org/10.1016/j.neuroimage.2014.06.008.
Gamberini, M., Passarelli, L., Filippini, M., Fattori, P., Galletti, C., 2021. Vision for action: thalamic and cortical inputs to the macaque superior parietal lobule. Brain Struct. Funct. 226 (9), 2951–2966. https://doi.org/10.1007/s00429-021-02377-7.
Geirhos, R., Medina Temme, C.R., Rauber, J., Schütt, H.H., Bethge, M., Wichmann, F.A., 2018. Generalisation in humans and deep neural networks. https://github.com/rgeirhos/generalisation-humans-DNNs.
Georgieva, S., Peeters, R., Kolster, H., Todd, J.T., Orban, G.A., 2009. The Processing of Three-Dimensional Shape from Disparity in the Human Brain 29 (3), 727–742. https://doi.org/10.1523/JNEUROSCI.4753-08.2009.
Georgieva, S., Todd, J., Peeters, R., Orban, G.A., 2008. The extraction of 3D shape from texture and shading in the human brain. Cerebral Cortex October 18, 2416–2438. https://doi.org/10.1093/cercor/bhn002.
Gillebert, R., Schaeverbeke, J., Bastin, C., Neyens, V., Bruffaerts, X., Weer, A. De, et al., 2015. 3D Shape Perception in Posterior Cortical Atrophy : Visual l (37), 12673–12692. https://doi.org/10.1523/JNEUROSCI.3651-14.2015.
Goldstein-Marcusohn, Y., Asaad, R., Asaad, L., Freud, E., 2024. The large-scale organization of shape processing in the ventral and dorsal pathways is dissociable from attention. Cerebr. Cortex 34 (6). https://doi.org/10.1093/cercor/bhae221.
Goodale, M.A., Milner, A.D., 1992. Separate visual pathways for perception and action. Trends Neurosci. 11 (1), 80–92. https://doi.org/10.1093/litthe/11.1.80.
Goodale, M.A., Milner, A.D., Melvyn, A., 1992. Separate visual pathways for perception and action. Trends Neurosci. 15 (1), 20–25.
Grill-Spector, K., Kourtzi, Z., Kanwisher, N., 2001. The lateral occipital complex and its role in object recognition. Vis. Res. 41 (10–11), 1409–1422.
Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., Malach, R., 1998. Cue-invariant activation in object-related areas of the human occipital lobe. Neuron 21 (1), 191–202. https://doi.org/10.1016/S0896-6273(00)80526-7.
Grill-Spector, K., Weiner, K.S., 2014. The functional architecture of the ventral temporal cortex and its role in categorization. Nat. Rev. Neurosci. 15 (8), 536–548. https://doi.org/10.1038/nrn3747.
Güçlü, U., van Gerven, M.A.J., 2015. Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. J. Neurosci. 35 (27), 10005–10014. https://doi.org/10.1523/JNEUROSCI.5023-14.2015.
Hayashi, T., Umeda, C., Cook, N.D., 2007. An fMRI study of the reverse perspective illusion 63. https://doi.org/10.1016/j.brainres.2007.05.073.
Hinkle, D.A., Connor, C.E., 2002. Three-dimensional orientation tuning in macaque area V4. Nat. Neurosci. 5 (7), 665–670. https://doi.org/10.1038/nn875.
Hung, C.-C., Carlson, E.T., Connor, C.E., 2012. Medial axis shape coding in macaque inferotemporal cortex. Neuron 74 (6), 1099–1113. https://doi.org/10.1016/j.neuron.2012.04.029.
Iwaki, S., Bonmassar, G., Belliveau, J.W., 2013. Dynamic cortical activity during the perception of three-dimensional object shape from two-dimensional random-dot motion. J. Integr. Neurosci. 12 (3). https://doi.org/10.1142/S0219635213500210.Dynamic.
Janssen, P., Scherberger, H., 2015. Visual guidance in control of grasping. Annu. Rev. Neurosci. 38, 69–86. https://doi.org/10.1146/annurev-neuro-071714-034028.
Jastorff, J., Orban, G.A., 2009. Human Functional Magnetic Resonance Imaging Reveals Separation and Integration of Shape and Motion Cues in Biological Motion Processing 29 (22), 7315–7329. https://doi.org/10.1523/JNEUROSCI.4870-08.2009.
Jeong, S.K., Xu, Y., 2016. Behaviorally relevant abstract object identity representation in the Human parietal cortex. J. Neurosci. 36 (5), 1607–1619. https://doi.org/10.1523/jneurosci.1016-15.2016.
Jolicoeur, P., 1985. The time to name disoriented natural objects. Mem. Cognit. 13 (4), 289–303.
Kar, K., Kubilius, J., Schmidt, K., Issa, E.B., DiCarlo, J.J., 2019. Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior. Nat. Neurosci. 22 (6), 974–983. https://doi.org/10.1038/s41593-019-0392-5.
Kim, T., Bair, W., Pasupathy, A., 2019. Neural coding for shape and texture in macaque area V4. J. Neurosci. 3073. https://doi.org/10.1523/JNEUROSCI.3073-18.2019, 18.
Konkle, T., Oliva, A., 2012. A real-world size Organization of object responses in occipitotemporal cortex. Neuron 74 (6), 1114–1124. https://doi.org/10.1016/j.neuron.2012.04.036.
Kubilius, J., Schrimpf, M., Nayebi, A., Bear, D., Yamins, D.L.K., Dicarlo, J.J., 2018. CORnet : modeling the neural mechanisms of core object recognition. BioRxiv Preprint 1–9.
Leek, E.C., Reppa, I., Arguin, M., 2005. The structure of three-dimensional object representations in human vision: evidence from whole-part matching. J. Exp. Psychol. Hum. Percept. Perform. 31 (4), 668–684. https://doi.org/10.1037/0096-1523.31.4.668.
Lescroart, M.D., Biederman, I., 2013. Cortical representation of medial axis structure. Cerebr. Cortex 23 (3), 629–637. https://doi.org/10.1093/cercor/bhs046.
Lescroart, M.D., Gallant, J.L., 2019. Human scene-selective areas represent 3D configurations of surfaces. Neuron 101 (1), 178–192.e7. https://doi.org/10.1016/j.neuron.2018.11.004.
Liu, C.H., Collin, C.A., Farivar, R., Chaudhuri, A., 2005. Recognizing faces defined by texture gradients. Percept. Psychophys. 67 (1), 158–167. https://doi.org/10.3758/BF03195019.
Logothetis, N.K., Pauls, J., 1995. Psychophysical and physiological evidence for viewer-centered object representations in the primate. Cerebr. Cortex 5 (3), 270–288. https://doi.org/10.1093/cercor/5.3.270.
Logothetis, N.K., Pauls, J., Bülthoff, H.H., Poggio, T., 1994. View-dependent object recognition by monkeys. Curr. Biol. 4 (5), 401–414. https://doi.org/10.1016/S0960-9822(00)00089-0.
Long, B., Yu, C.P., Konkle, T., 2018. Mid-level visual features underlie the high-level categorical organization of the ventral stream. Proc. Natl. Acad. Sci. U. S. A. 115 (38), E9015–E9024. https://doi.org/10.1073/pnas.1719616115.
Marcus, D.S., Harwell, J., Olsen, T., Hodge, M., Glasser, M.F., Prior, F., Marcus, D.S., 2011. Informatics and data mining tools and strategies for the Human Connectome Project 5 (June), 1–12. https://doi.org/10.3389/fninf.2011.00004.
Milner, A.D., Perrett, D.I., Johnston, R.S., Benson, P.J., Jordan, T.R., Heeley, D.W., et al., 1991. Perception and action in "visual form agnosia.". Brain 114 (1), 405–428. https://doi.org/10.1093/brain/114.1.405.
Mishkin, M., Ungerleider, L., Macko, K., 1983. Object vision and spatial vision two cortical pathways. Trends Neurosci.
Mocz, V., Jeong, S.K., Chun, M., Xu, Y., 2023. Multiple visual objects are represented differently in the human brain and convolutional neural networks. Sci. Rep. 13 (1), 1–16. https://doi.org/10.1038/s41598-023-36029-z.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Group, T.P., 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses : The PRISMA Statement 6 (7). https://doi.org/10.1371/journal.pmed.1000097.
Müller, V.I., Cieslik, E.C., Laird, A.R., Fox, P.T., Radua, J., Mataix-cols, D., et al., 2018. Ten simple rules for neuroimaging meta-analysis. Neurosci. Biobehav. Rev. 84, 151–161. https://doi.org/10.1016/j.neubiorev.2017.11.012. November 2017.
Mur, M., Meys, M., Bodurka, J., Goebel, R., Bandettini, P.A., Kriegeskorte, N., 2013. Human object-similarity judgments reflect and transcend the primate-IT object representation. Front. Psychol. 4 (MAR), 1–22. https://doi.org/10.3389/fpsyg.2013.00128.
Murphy, A.P., Ban, H., Welchman, A.E., 2013. Integration of texture and disparity cues to surface slant in dorsal visual cortex. J. Neurophysiol. 110 (1), 190–203. https://doi.org/10.1152/jn.01055.2012.
Naganuma, T., Nose, I., Inoue, K., Takemoto, A., 2005. Information Processing of Geometrical Features of a Surface Based on Binocular Disparity Cues : an Fmri Study, vol. 51, pp. 147–155. https://doi.org/10.1016/j.neures.2004.10.009.
Nishida, Y., Hayashi, C.A.O., Iwami, T., Kimura, M., Kani, K., Ito, R., et al., 2001. Stereopsis-processing regions in the human parieto-occipital cortex, 12 (10), 2259–2263.
Ogawa, A., Bordier, C., Macaluso, E., 2013. Audio-Visual Perception of 3D Cinematography : An fMRI Study Using Condition-Based and Computation-Based Analyses 8 (10). https://doi.org/10.1371/journal.pone.0076003.
Orban, G.A., Sunaert, S., Todd, J.T., Hecke, P. Van, Marchal, G., Universiteit, K., Gasthuisberg, U.Z., 1999. Human Cortical Regions Involved in Extracting Depth from Motion, vol. 24, pp. 929–940.
Paradis, A.L., Droulez, J., 2000. Visual Perception of Motion and 3-D Structure from Motion : an Fmri Study, pp. 772–783.

Pasupathy, A., Kim, T., Popovkina, D.V., 2019. Object shape and surface properties are jointly encoded in mid-level ventral visual cortex. Curr. Opin. Neurobiol. 58, 199–208. https://doi.org/10.1016/j.conb.2019.09.009.

Peissig, J.J., Tarr, M.J., 2006. Visual object recognition: do we know more now than we did 20 years ago? Annu. Rev. Psychol. 0 (0). https://doi.org/10.1146/annurev.psych.57.102904.190114, 060822030849013.

Pelli, D.G., Majaj, N.J., Raizman, N., Christian, C.J., Kim, E., Palomares, M.C., 2009. Cognitive neuropsychology Grouping in object recognition: the role of a Gestalt law in letter identification. Cogn. Neuropsychol. 26 (1), 36–49. https://doi.org/10.1080/13546800802550134.

Peuskens, H., Claeys, K.G., Todd, J.T., Norman, J.F., Van Hecke, P., Orban, G.A., 2004. Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. J. Cognit. Neurosci. 16 (4), 665–682. https://doi.org/10.1162/089892904323057371.

Poggio, T., Edelman, S., 1990. A network that learns to recognize three-dimensional objects. Nature 343 (6255), 263–266. https://doi.org/10.1038/343263a0.

Poggio, T., Riesenhuber, M., 1999. Hierarchical models of object recognition in cortex. Nat. Neurosci. 2 (11), 1019–1025. Retrieved from. http://www.nature.com/doifinder/10.1038/14819%5Cnpapers3://publication/doi/10.1038/14819.

Rajalingham, R., Schmidt, K., DiCarlo, J.J., 2015. Comparison of object recognition behavior in human and monkey. J. Neurosci. 35 (35), 12127–12136. https://doi.org/10.1523/jneurosci.0573-15.2015.

Riesenhuber, M., Poggio, T., 1997. Just one view: invariances in inferotemporal cell tuning. Adv. Neural Inf. Process. Syst. 137–144.

Riesenhuber, M., Poggio, T., 2000. Models of object recognition. Nat. Neurosci. 3, 1199–1204.

Salimi-Khorshidi, G., Smith, S.M., Keltner, J.R., Wager, T.D., Nichols, T.E., 2009a. Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. Neuroimage 45 (3), 810–823. https://doi.org/10.1016/J.NEUROIMAGE.2008.12.039.

Salimi-Khorshidi, G., Smith, S.M., Keltner, J.R., Wager, T.D., Nichols, T.E., 2009b. Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. Neuroimage 45 (3), 810–823. https://doi.org/10.1016/J.NEUROIMAGE.2008.12.039.

Saygin, A.P., Wilson, S.M., Jr, D.J.H., Bates, E., Sereno, M.I., 2004. Point-Light Biological Motion Perception Activates Human Premotor Cortex 24 (27), 6181–6188. https://doi.org/10.1523/JNEUROSCI.0504-04.2004.

Schindler, A., Bartels, A., 2016. NeuroImage motion parallax links visual motion areas and scene regions. Neuroimage 125, 803–812. https://doi.org/10.1016/j.neuroimage.2015.10.066.

Sereno, M.E., Trinath, T., Augath, M., Logothetis, N.K., 2002. Three-dimensional shape representation in Monkey Cortex. Neuron 33. Retrieved from Janssen website: www.neuron.org/cgi/content/full/33/4/635/DC1.

Snow, J.C., Pettypiece, C.E., Mcadam, T.D., Mclean, A.D., Stroman, P.W., Goodale, M.A., Culham, J.C., 2011. Bringing the real world into the fMRI scanner: repetition effects for pictures versus real objects. https://doi.org/10.1038/srep00130.

Srinath, R., Emonds, A., Wang, Q., Lempel, A.A., Dunn-Weiss, E., Connor, C.E., Nielsen, K.J., 2021. Early emergence of solid shape coding in natural and deep network vision. Curr. Biol. 31 (1), 51–65.e5. https://doi.org/10.1016/j.cub.2020.09.076.

Su, J., Vargas, D.V., Sakurai, K., 2019. One pixel attack for fooling deep neural networks. IEEE Trans. Evol. Comput. 1. https://doi.org/10.1109/tevc.2019.2890858, 1.

Suchan, B., Botko, R., Gizewski, E., Forsting, M., Daum, I., 2006. Neural Substrates of Manipulation in Visuospatial Working Memory, vol. 139, pp. 351–357. https://doi.org/10.1016/j.neuroscience.2005.08.020.

Takemura, H., Rokem, A., Winawer, J., Yeatman, J.D., Wandell, B.A., Pestilli, F., 2015. A major human white matter pathway between dorsal and ventral visual Cortex. Cerebr. Cortex 26 (5), 2205–2214. https://doi.org/10.1093/cercor/bhv064.

Tarr, M.J., Pinker, S., 1989. Mental rotation and orientation-dependence shape recognition projections when it is seen at different orientations , sizes , and positions ? Formed to the size , orientation , and location of the " nearest " one . the effect of orientation on the amount o. Cogn. Psychol. 21, 233–282.

Tench, C.R., Tanasescu, R., Constantinescu, C.S., Cottam, W.J., Auer, D.P., 2020. Coordinate based meta-analysis of networks in neuroimaging studies. Neuroimage 205, 116259. https://doi.org/10.1016/J.NEUROIMAGE.2019.116259.

Tsao, D.Y., Vanduffel, W., Sasaki, Y., Fize, D., Knutsen, T.A., Mandeville, J.B., et al., 2003. Stereopsis Activates V3A and Caudal Intraparietal Areas in Macaques and Humans, vol. 39, pp. 555–568, 1967.

Uji, M., Lingnau, A., Cavin, I., Vishwanath, D., Hibbard, P.B., 2019. Identifying Cortical Substrates Underlying the Phenomenology of Stereopsis and Realness : A Pilot fMRI Study 13 (July), 1–14. https://doi.org/10.3389/fnins.2019.00646.

Ullman, S., Assif, L., Fetaya, E., Harari, D., 2016. Atoms of recognition in human and computer vision. Proc. Natl. Acad. Sci. U. S. A. 113 (10), 2744–2749. https://doi.org/10.1073/pnas.1513198113.

Ullman, S., Basri, R., 1991. Recognition by linear combinations of models. IEEE Trans. Pattern Anal. Mach. Intell. 13 (10), 992–1006.

Vaziri, S., Carlson, E.T., Wang, Z., Connor, C.E., 2014. A channel for 3D environmental shape in anterior inferotemporal cortex. Neuron 84 (1), 55–62. https://doi.org/10.1016/j.neuron.2014.08.043.

Verhoef, B.E., Vogels, R., Janssen, P., 2016. Binocular depth processing in the ventral visual pathway. Phil. Trans. Biol. Sci. 371 (1697). https://doi.org/10.1098/rstb.2015.0259.

Wang, L., Mruczek, R.E.B., Arcaro, M.J., Kastner, S., 2015. Probabilistic Maps of Visual Topography in Human Cortex, pp. 3911–3931. https://doi.org/10.1093/cercor/bhu277. October.

Yamamoto, T., Takahashi, S., Hanakawa, T., Urayama, S., Aso, T., Fukuyama, H., Ejima, Y., 2008. Neural Correlates of the Stereokinetic Effect Revealed by Functional Magnetic Resonance Imaging Tetsuya Yamamoto, vol. 8, pp. 1–17. https://doi.org/10.1167/8.10.14.

Yamane, Y., Carlson, E.T., Bowman, K.C., Wang, Z., Connor, E., 2008. A neural code for three-dimensional object shape in macaque inferotemporal cortex. Nat. Neurosci. 11 (11), 1352–1360. https://doi.org/10.1038/nn.2202.A.

Yamins, D.L.K., Hong, H., Cadieu, C.F., Solomon, E.A., Seibert, D., DiCarlo, J.J., 2014. Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proc. Natl. Acad. Sci. U. S. A. 111 (23), 8619–8624. https://doi.org/10.1073/pnas.1403112111.