

Moving and Static Faces, Bodies, Objects, and Scenes Are Differentially Represented across the Three Visual Pathways

Emel Küçük¹, Matthew Foxwell¹, Daniel Kaiser^{1,2,3} , and David Pitcher¹ 

Abstract

■ Models of human cortex propose the existence of neuroanatomical pathways specialized for different behavioral functions. These pathways include a ventral pathway for object recognition, a dorsal pathway for performing visually guided physical actions, and a recently proposed third pathway for social perception. In the current study, we tested the hypothesis that different categories of moving stimuli are differentially processed across the dorsal and third pathways according to their behavioral implications. Human participants ($n = 30$) were scanned with fMRI while viewing moving and static stimuli from four categories (faces, bodies, scenes, and objects). A whole-brain

group analysis showed that moving bodies and moving objects increased neural responses in the bilateral posterior parietal cortex, parts of the dorsal pathway. By contrast, moving faces and moving bodies increased neural responses, the superior temporal sulcus, part of the third pathway. This pattern of results was also supported by a separate ROI analysis showing that moving stimuli produced more robust neural responses for all visual object categories, particularly in lateral and dorsal brain areas. Our results suggest that dynamic naturalistic stimuli from different categories are routed in specific visual pathways that process dissociable behavioral functions. ■

INTRODUCTION

Explaining the neural processes that enable humans to perceive, understand, and interact with the people, places, and objects we encounter in the world is a fundamental aim of visual neuroscience. An experimentally rich theoretical approach in pursuit of this goal has been to show that dissociable cognitive functions are performed in anatomically segregated cortical pathways. For example, influential models of the visual cortex propose it contains two functionally distinct pathways: a ventral pathway specialized for visual object recognition, and a dorsal pathway specialized for performing visually guided physical actions (Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013; Kravitz, Saleem, Baker, & Mishkin, 2011; Milner & Goodale, 1995; Ungerleider & Mishkin, 1982). Despite the influence these models, neither can account for the neural processes that underpin human social interaction. Social interactions are predicated on visually analyzing and understanding the actions of others and responding appropriately. One region of the brain in particular, the superior temporal sulcus (STS), computes the sensory information that facilitates these processes (Kilner, 2011; Allison, Puce, & McCarthy, 2000; Perrett, Hietanen, Oram, & Benson, 1992). We recently proposed the existence of a visual pathway specialized for social perception (Pitcher,

2021; Pitcher & Ungerleider, 2021). This pathway projects from the primary visual cortex into the STS, via the motion-selective area V5/MT (Watson et al., 1993). The aim of the current study was to test a prediction of our model by contrasting the response to moving and static visual stimuli across the proposed three visual pathways. Specifically, we predicted moving biological stimuli (e.g., faces and bodies) are preferentially processed along a dedicated neural pathway that includes V5/MT and the STS, compared with moving stimuli of nonbiological categories.

The STS selectively responds to moving biological stimuli (e.g., faces and bodies) and computes the visual social cues that help us understand and interpret the actions of other people. These include facial expressions (Sliwinski, Elson, & Pitcher, 2020; Phillips et al., 1997), eye gaze (Pourtois et al., 2004; Puce, Allison, Bentin, Gore, & McCarthy, 1998; Campbell, Heywood, Cowey, Regard, & Landis, 1990), body movements (Beauchamp, Lee, Haxby, & Martin, 2003; Grossman & Blake, 2002), and the audio-visual integration of speech (Beauchamp, Nath, & Pasalar, 2010; Calvert et al., 1997). However, the connectivity between early visual cortex and the STS remains poorly characterized. This led some researchers to view the STS as an extension of the ventral pathway, rather than as a functionally and anatomically independent pathway in its own right. For example, models of face processing propose that all facial aspects (e.g., identity and expression

¹University of York, ²Justus-Liebig-Universität Gießen, ³Philipps-Universität Marburg and Justus-Liebig-Universität Gießen

recognition) are processed using the same early visual mechanisms (Pitcher, Walsh, & Duchaine, 2011; Calder & Young, 2005; Haxby, Hoffman, & Gobbini, 2000; Bruce & Young, 1986) before diverging at higher levels of processing, rather than as dissociable processes that begin in early visual cortex. Contrary to this view, alternate models propose that dynamic facial information is preferentially processed in a dissociable cortical pathway that projects from early visual cortex, via the motion-selective area V5/MT directly into the STS (Duchaine & Yovel, 2015; Pitcher, Duchaine, & Walsh, 2014; LaBar, Crupain, Voyvodic, & McCarthy, 2003; O'Toole, Roark, & Abdi, 2002). This is consistent with our model of the third visual pathway that predicts that moving faces and bodies will selectively evoke neural activity in a pathway projecting from V1 to the STS, via V5/MT as shown in Figure 1 (Pitcher & Ungerleider, 2021).

The cognitive and behavioral functions performed in a particular brain area can be deduced (at least partially) by the anatomical connectivity of that area (Pitcher et al., 2020; Kravitz et al., 2011, 2013; Boussaoud, Ungerleider, & Desimone, 1990; Desimone & Ungerleider, 1986; Ungerleider & Desimone, 1986). This approach leads to hierarchical models that dissociate cognitive functions based on behavioral goals (e.g., visually recognizing a friend, or reaching to shake their hand, or interpreting their mood). An alternate conceptual approach for studying the cognitive functionality of the brain has been to take a modular approach (Fodor, 1983). Modularity favors the view that cortex contains discrete cortical patches that respond to specific visual characteristics such as motion (Watson et al., 1993), or to stimulus categories including objects (Malach et al., 1995), faces (Kanwisher, McDermott, & Chun, 1997), scenes (Epstein & Kanwisher, 1998), and bodies (Downing, Jiang, Shuman, & Kanwisher, 2001). Although there has sometimes been an inherent tension between anatomical and modular models of

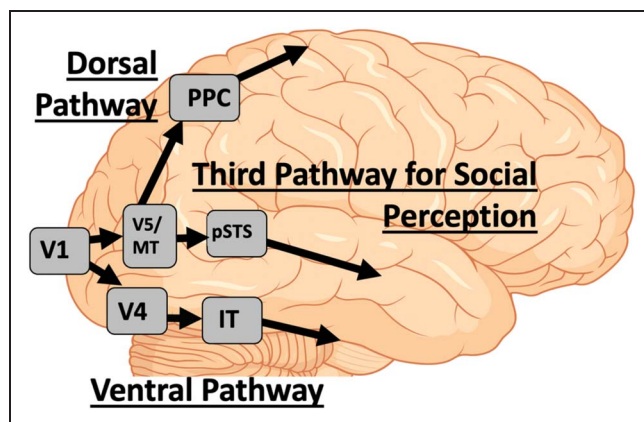


Figure 1. The three visual pathways (Pitcher & Ungerleider, 2021). The ventral pathway projects from V1 via ventral V4 into the IT and anterior IT cortex. The dorsal pathway projects from V1 to V5/MT into the PPC and then to the motor cortex. The third visual pathway for social perception projects from V1 to V5/MT and then to the pSTS.

cortical organization (Kanwisher, 2010; Hein & Knight, 2008), it has also been argued that different conceptual and methodical approaches can reveal cortical functionality at different levels of understanding (de Haan & Cowey, 2011; Grill-Spector & Malach, 2004; Walsh & Butler, 1996).

The aim of the present study was to measure the neural responses at both the hierarchical and modular level by manipulating the visual stimuli on two dimensions: moving versus static stimuli, or the object category of the stimuli (faces, bodies, scenes, and objects). Participants were scanned using fMRI while viewing 3-sec videos or static images taken from the videos. The visual categories included were faces, bodies, objects, scenes, and scrambled objects (Figure 2). This set of stimuli can be used to identify a series of category-selective areas across the brain that include face areas (Haxby et al., 2000), body areas (Peelen & Downing, 2007), scene areas (Epstein, 2008), and object areas (Malach et al., 1995). We have previously used this design to functionally dissociate the neural response across face areas (Pitcher et al., 2014) and across the lateral and ventral surfaces of the occipitotemporal cortex (Pitcher, Ianni, & Ungerleider, 2019). However, these prior studies lacked the necessary experimental conditions and the whole-brain coverage to systematically compare the response to motion and visual category across the entire brain. The present study systematically compares the responses to different moving and static stimulus categories across the whole brain, as well as in targeted ROI analyses. This provides detailed insights into how dynamically presented visual categories are routed along the three visual pathways.

METHODS

Participants

Thirty participants (20 female; age range = 18–48 years; mean age = 23 years) with normal, or corrected-to-normal, vision gave informed consent as directed by the ethics committee at the University of York. The sample size was based on convenience sampling and chosen to exceed $n = 27$ (offering 80% power for detecting medium effects of $d = 0.5$ in one-sided t tests) and prior studies that used these stimuli as localizers (Sliwinska et al., 2022; Handwerker et al., 2020). Data from 24 participants was collected for a previous fMRI experiment (Nikel, Sliwinska, Kucuk, Ungerleider, & Pitcher, 2022) and reanalyzed for the current study.

Stimuli

Dynamic stimuli were 3-sec movie clips of faces, bodies, scenes, objects, and scrambled objects designed to localize the category-selective brain areas of interest (Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011). There were 60 movie clips for each category in which distinct

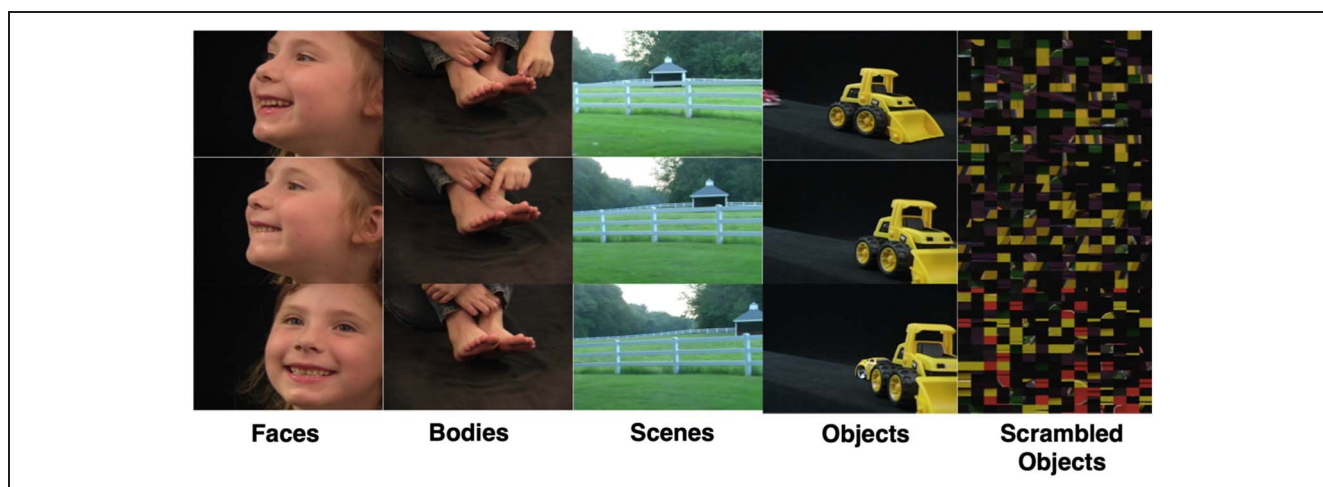


Figure 2. Examples of the static images taken from the 3-sec movie clips depicting faces, bodies, scenes, objects, and scrambled objects. Still images taken from the beginning, middle, and end of the corresponding movie clip.

exemplars appeared multiple times. Movies of faces and bodies were filmed on a black background and framed close-up to reveal only the faces or bodies of seven children as they danced or played with toys or adults (who were out of frame). Fifteen different locations were used for the scene stimuli, which were mostly pastoral scenes shot from a car window while driving slowly through leafy suburbs, along with some other films taken while flying through canyons or walking through tunnels that were included for variety. Fifteen different moving objects were selected that minimized any suggestion of animacy of the object itself or of a hidden actor pushing the object (these included mobiles, windup toys, toy planes and tractors, balls rolling down sloped inclines). Scrambled objects were constructed by dividing each object movie clip into a 15×15 box grid and spatially rearranging the location of each of the resulting movie frames. Within each block, stimuli were randomly selected from within the entire set for that stimulus category (faces, bodies, scenes, objects, scrambled objects). This meant that the same actor, scene, or object could appear within the same block but, given the number of stimuli, this did not occur regularly.

Static stimuli were identical in design to the dynamic stimuli except that, in place of each 3-sec movie, we presented three different still images taken from the beginning, middle, and end of the corresponding movie clip. Each image was presented for 1 sec with no ISI, to equate the total presentation time with the corresponding dynamic movie clip (Figure 2). This same stimulus set has been used in our prior fMRI studies of category-selective areas (Pitcher, Sliwinska, & Kaiser, 2023; Sliwinska et al., 2022).

Procedure and Data Acquisition

Functional data were acquired over eight blocked-design functional runs lasting 234 sec each. Each functional run

contained three 18-sec rest blocks, at the beginning, middle, and end of the run, during which a series of six uniform color fields were presented for 3 sec. The colors were included because the stimuli were originally designed for scanning children, and it was thought this would make the session more engaging. Participants were instructed to watch the movies and static images but were not asked to perform any overt task.

Functional runs presented either movie clips (four dynamic runs) or sets of static images taken from the same movies (four static runs). For the dynamic runs, each 18-sec block contained six 3-sec movie clips from that category. For the static runs, each 18-sec block contained eighteen 1-sec still snapshots, composed of six triplets of snapshots taken at 1-sec intervals from the same movie clip. Dynamic/static runs were run in the following order: two dynamic, two static, two dynamic, two static. Ordering of the stimulus blocks for each category within each run was pseudorandomized within and across participants.

Imaging data were acquired using a 3 T Siemens Magnetom Prisma MRI scanner (Siemens Healthcare) at the University of York. Functional images were acquired with a 20-channel phased array head coil and a gradient-echo EPI sequence (38 interleaved slices, repetition time (TR) = 3 sec, echo time (TE) = 30 msec, flip angle = 90° ; voxel size = 3 mm isotropic; matrix size = 128×128) providing whole-brain coverage. Slices were aligned with the anterior to posterior commissure line. Structural images were acquired using the same head coil and a high-resolution, T1-weighted, 3-D, fast spoiled gradient (spoiled gradient recall) sequence (176 interleaved slices, TR = 7.8 sec, TE = 3 msec, flip angle = 20° ; voxel size = 1 mm isotropic; matrix size = 256×256).

Imaging Analysis

fMRI data were analyzed using AFNI (<https://afni.nimh.nih.gov/afni>). Images were slice-time corrected and realigned

to the third volume of the first functional run and to the corresponding anatomical scan. All data were motion corrected, and any TRs in which a participant moved more than 0.3 mm in relation to the previous TR were discarded from further analysis. The volume-registered data were spatially smoothed with a 4-mm FWHM Gaussian kernel. Signal intensity was normalized to the mean signal value within each run and multiplied by 100 so that the data represented percent signal change from the mean signal value before analysis.

Data from all runs were entered into a general linear model by convolving a gamma hemodynamic response function with the regressors of interest (faces, bodies, scenes, objects, and scrambled objects) for dynamic and static functional runs. Regressors of no interest (e.g., six head movement parameters obtained during volume registration and AFNI's baseline estimates) were also included in the general linear model. Data from all 30 participants were entered into a group whole-brain, two-way, mixed-effects ANOVA with Motion (moving, static) and Visual Category (faces, bodies, objects, scenes, and scrambled objects) as fixed effects and participants as the random effect. The interaction effects obtained in this ANOVA are shown in Figure 3. We calculated significant clusters based on an uncorrected threshold of $p < .001$ at the voxel level and $p < .05$ at the cluster level (calculated using 3dClustSim). For display purposes, volumetric whole-brain maps were projected onto an inflated cortical surface using SUMA (<https://afni.nimh.nih.gov/Suma>).

To understand what stimulus categories were driving the interaction, we calculated the main effects for each of the four visual categories (faces, bodies, scenes, and objects) for the moving (Figure 4) and static (Figure 5) stimulus conditions. Activation maps show clusters above the statistical threshold $p < .001$ at the voxel level and $p < .05$ at the cluster level.

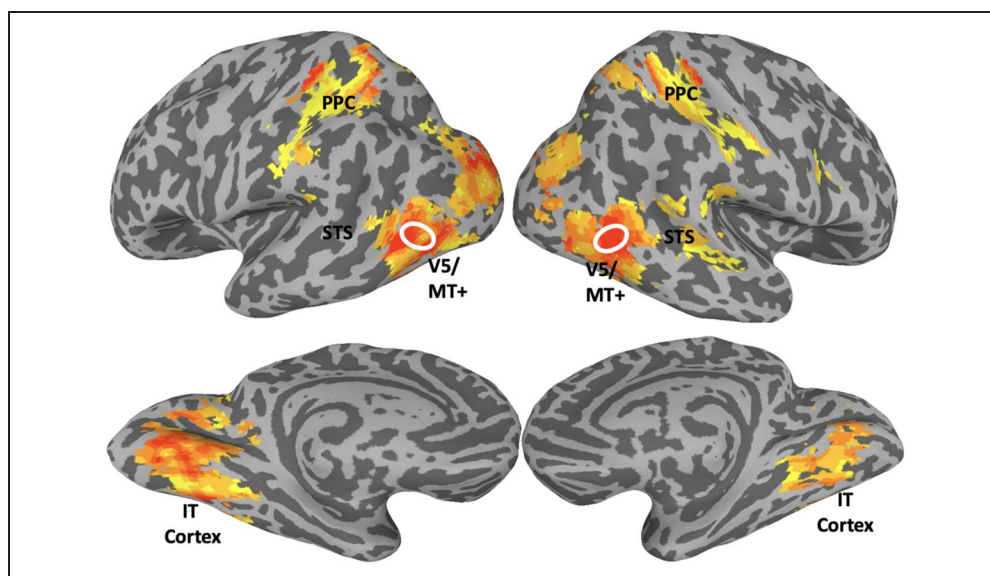
.05 at the cluster level. Finally, we calculated four group whole-brain contrasts showing moving greater than static neural activity separately for each visual category (faces, bodies, objects, and scenes). Activation maps for each contrast are shown in Figure 6 (again calculated using a statistical threshold $p < .001$ at the voxel level and $p < .05$ at the cluster level).

RESULTS

Data from all 30 participants were entered into a group whole-brain, mixed-effects, two-way ANOVA to establish how moving and static stimuli were differentially processed across the three visual pathways. The results of the 2 (Motion: moving, static) \times 5 (Category: faces, bodies, objects and scenes) interaction obtained in this analysis are shown in Figure 3. Significant interactions were apparent across parts of all three visual pathways. These included inferior temporal (IT) cortex in the ventral pathway, the posterior parietal cortex (PPC) in the dorsal pathway, and the STS in the third pathway.

To unpack which stimuli were driving the interactions between motion and category, we plotted the main effects for each category (faces, bodies, objects, and scenes) for the moving (Figure 4) and static (Figure 5) conditions. For the moving face (Figure 4A) and moving body (Figure 4B), we observed neural activity that extended from visual cortex into the STS, this was bilateral for moving bodies but restricted to the right STS for moving faces. This is consistent with the functional role of the third pathway for social perception, and the laterality of face processing in the right STS was also shown in our prior TMS study (Sliwinski & Pitcher, 2018). We also observed significant neural activity in the bilateral PPC for moving body and moving object stimuli, part of the vision for action pathway in the dorsal pathway (Milner & Goodale, 1995). We observed

Figure 3. Interaction between motion and visual categories obtained in the group whole-brain, mixed-effects ANOVA. The activation map shows the results of the two-way ANOVA between Motion (moving, static) and Category (faces, bodies, objects, scenes, and scrambled). Significant interactions were calculated using an uncorrected threshold of $p < .001$ at the voxel level and $p < .05$ at the cluster level. Significant clusters can be observed across the three visual pathways. These include bilateral IT cortex in the ventral pathway, bilateral PPC in the dorsal pathway, and the right STS in the third pathway. The location of V5/MT is circled in white.



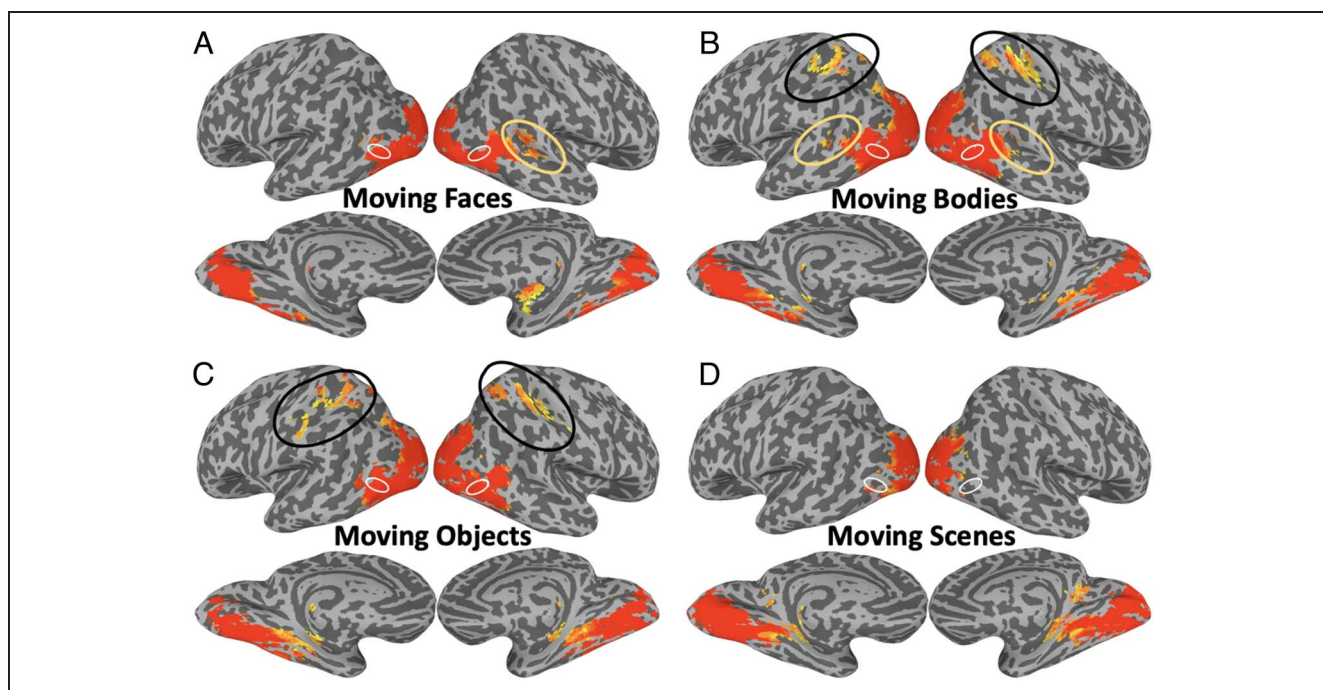


Figure 4. Activation maps showing the main effects of the group whole-brain mixed ANOVA for the four visual categories in the moving condition (uncorrected threshold of $p < .001$ at the voxel level and $p < .05$ at the cluster level). Figures show the activation maps for Moving Faces (A), Moving Bodies (B), Moving Objects (C), and Moving Scenes (D). Consistent with the established functional roles of the dorsal pathway, there were significant activations for moving bodies and moving objects in the PPC (circled in black). There were also significant activations for moving bodies in the bilateral STS and for moving faces in the right STS (circled in yellow). This is consistent with the functional role of the third pathway, which selectively processes moving biological stimuli. Neural activity for all four categories was observed in IT cortex, which is consistent with the visual recognition role computed in the ventral stream. The location of V5/MT is circled in white.

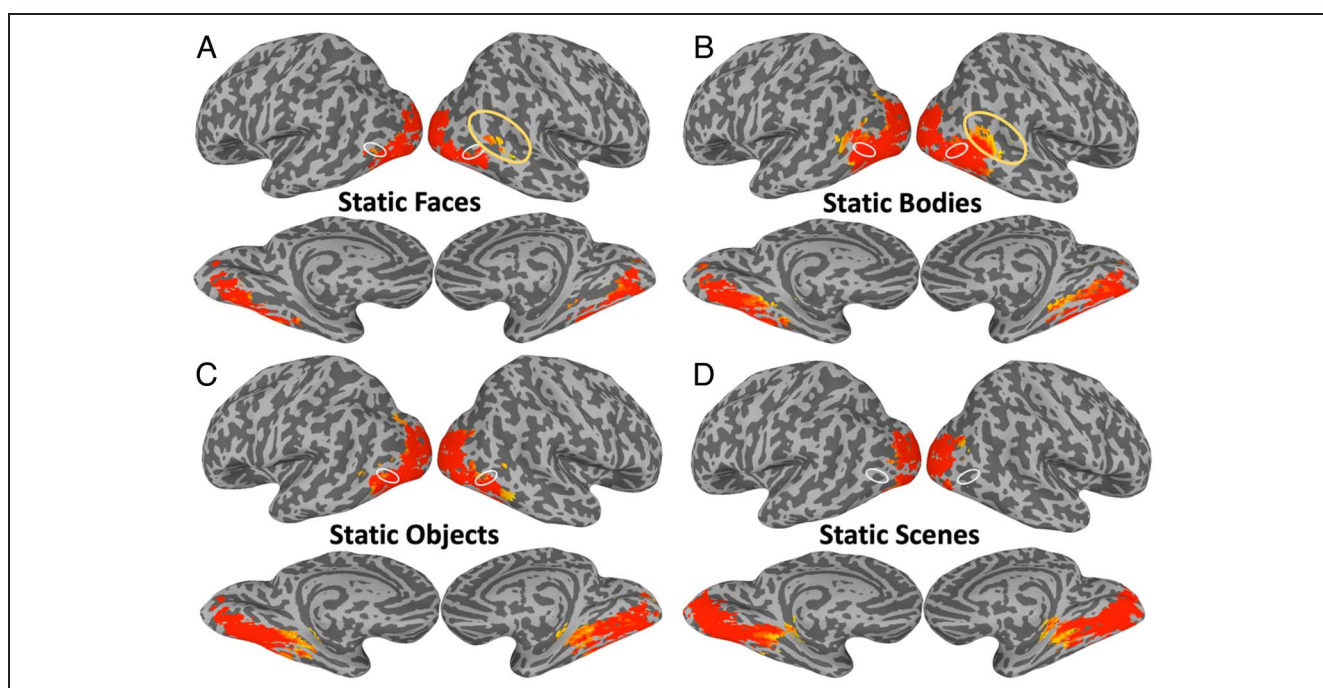


Figure 5. Activation maps showing the main effects of the group whole-brain mixed ANOVA for the four visual categories in the static condition (uncorrected threshold of $p < .001$ at the voxel level and $p < .05$ at the cluster level). Figures show the activation maps for Static Faces (A), Static Bodies (B), Static Objects (C), and Static Scenes (D). Neural activity for static faces and static bodies (circled in yellow) was observed in the right STS. Neural activity for all four categories was observed in IT cortex, which is consistent with the visual recognition function computed in the ventral pathway. Neural activity in the PPC and other parts of the dorsal pathway was not observed at this statistical threshold suggesting the dorsal pathway preferentially processes moving stimuli. The location of V5/MT is circled in white.

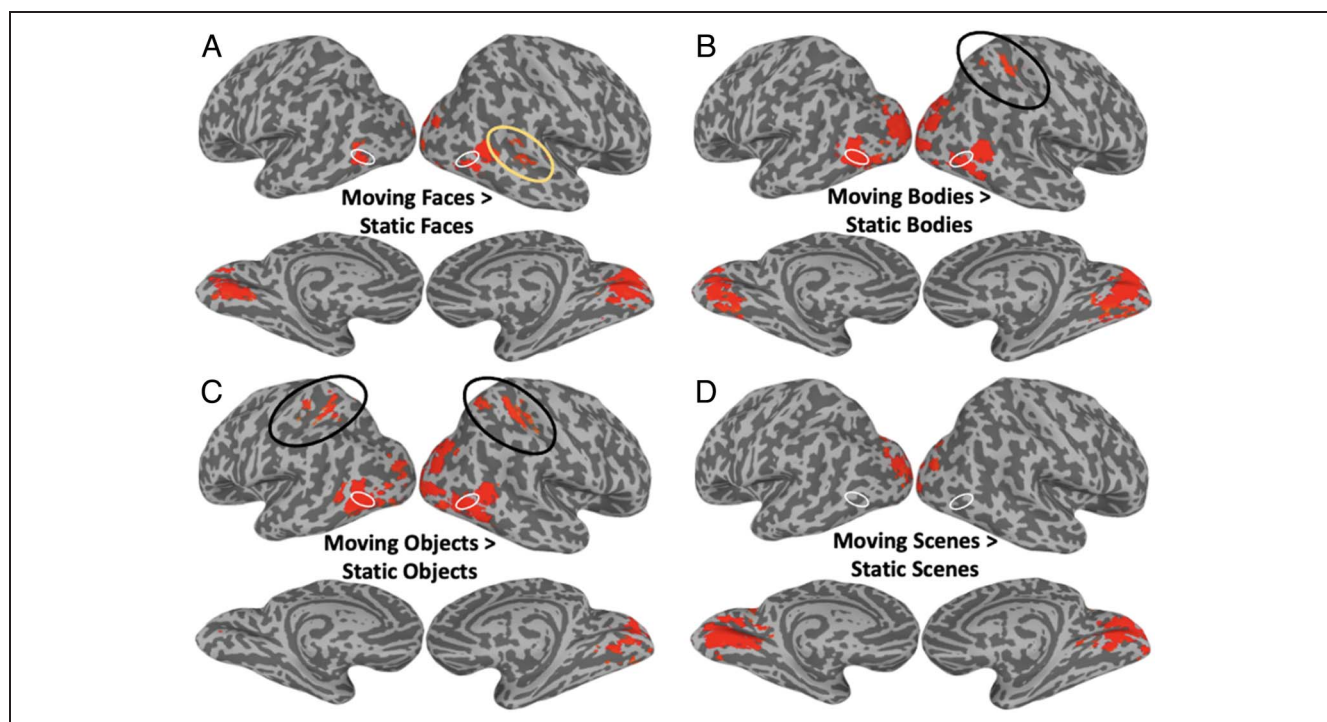


Figure 6. Results of four group whole-brain contrasts showing moving greater than static neural activity (uncorrected threshold of $p < .001$ at the voxel level and $p < .05$ at the cluster level) for each visual category separately (faces, bodies, objects, and scenes). Moving faces > static faces produced neural activity in the right STS (circled in yellow), part of the third pathway (A). Moving bodies > static bodies and moving objects > static objects produced neural activity in the right PPC (circled in black; B and C) and produced neural activity in the bilateral PPC (circled in black). The location of V5/MT is circled in white.

significant activity for all four categories in ventral IT cortex, part of the ventral pathway for visual recognition.

The main effect for static faces (Figure 5A) and for static bodies (Figure 5B) showed neural activity in the right STS. There was also significant activation for all four static categories in ventral IT cortex. By contrast, none of the four conditions revealed any significant activations in the parietal cortex. Finally, to verify the differential effects for moving stimuli across the three pathways, we contrasted the moving greater than static conditions for each visual category (Figure 6). These results were partially consistent with the main effects analysis showing selective activity for moving faces > static faces (Figure 6A) in the right STS, but there were no significant clusters in the STS for the moving body > static body contrast (Figure 6B). The moving objects > static objects contrast showed bilateral PPC activity (Figure 6C), whereas moving bodies > static bodies showed PPC activity in the right hemisphere only.

ROI Analysis

We also conducted a separate analysis to evaluate the use of moving stimuli for functionally localizing category-selective regions in the three visual pathways. We analyzed data for all participants individually to separately localize the ROIs using dynamic and static stimuli for all four stimulus categories (faces, scenes, bodies, and objects). Significance maps were calculated for each participant

individually, using an uncorrected statistical threshold of $p < .001$ for all four contrasts of interest. If the ROI was present in that participant, we identified the Montreal Neurological Institute (MNI) coordinates of the peak voxel and the number of contiguous voxels in the ROI. Mean results are displayed in Table 1.

Face-selective ROIs were identified in two separate analyses, using a contrast of moving faces greater than moving objects and a contrast of static faces greater than static objects. We attempted to localize face-selective voxels in five commonly studied face ROIs. These were the fusiform face area (FFA), the occipital face area, the posterior superior temporal sulcus (pSTS), the amygdala, and pFC. Results showed that face ROIs were present in more participants in the right hemisphere. The pSTS and pFC were identified in more participants when defined using moving than static stimuli. The pSTS and pFC ROIs were also larger when identified with moving stimuli. These results are consistent with some prior studies (Nikel et al., 2022; Pitcher et al., 2019; Fox, Iaria, & Barton, 2009), but it is important to note that other studies have demonstrated that the FFA can also exhibit a greater response to moving faces than static faces (Pilz, Vuong, Bühlhoff, & Thornton, 2011; Schultz & Pilz, 2009).

Scene-selective ROIs were identified in two separate analyses, using a contrast of moving scenes greater than moving objects and a contrast of static scenes greater than

Table 1. The Results of the ROI Analysis for Dynamic and Static Stimuli for Face, Scene, Body, and Object Areas in 30 Participants

			<i>Right Hemisphere</i>					<i>Left Hemisphere</i>				
			<i>MNI Coordinates</i>			<i>Mean Voxel #</i>	<i>ROI Present</i>	<i>MNI Coordinates</i>			<i>Mean Voxel #</i>	<i>ROI Present</i>
			<i>x</i>	<i>y</i>	<i>z</i>			<i>x</i>	<i>y</i>	<i>z</i>		
<u>Faces</u>	FFA	Dynamic	43	−52	−18	70	29/30	−42	−53	−19	60	29/30
		Static	42	−53	−18	63	27/30	−42	−55	−20	47	27/30
	OFA	Dynamic	40	−82	−9	64	29/30	−38	−81	−12	56	24/30
		Static	45	−80	−7	69	26/30	−45	−79	−8	52	20/30
	pSTS	Dynamic	53	−47	10	144	28/30	−55	−45	10	97	21/30
		Static	54	−51	9	56	20/30	−52	−49	8	23	17/30
	Amygdala	Dynamic	18	−4	−14	22	9/30	−19	−2	−15	18	7/30
		Static	23	−2	−15	12	3/30	−14	1	−16	27	1/30
	PFC	Dynamic	41	11	40	50	20/30	−48	11	33	42	8/30
		Static	44	11	40	27	8/30	−39	14	43	87	7/30
<u>Scenes</u>	PPA	Dynamic	23	−43	−12	N/A	29/30	−22	−47	−9	N/A	26/30
		Static	24	−42	−13	56	29/30	−23	−46	−9	42	25/30
	RSC	Dynamic	20	−56	12	N/A	29/30	−19	−59	13	N/A	19/30
		Static	22	−56	12	64	27/30	−20	−57	10	49	16/30
	OPA	Dynamic	43	−78	24	59	27/30	−39	−81	26	38	23/30
		Static	39	−79	24	34	21/30	−39	−83	26	31	20/30
<u>Bodies</u>	EBA	Dynamic	48	−70	5	243	30/30	−45	−73	10	203	29/30
		Static	53	−70	2	195	30/30	−51	−75	8	153	28/30
	FBA	Dynamic	39	−50	−17	45	23/30	−41	−44	−17	30	19/30
		Static	44	−50	−17	37	22/30	−41	−73	−18	31	17/30
<u>Objects</u>	LO	Dynamic	48	−73	−3	845	27/30	−47	−74	−1	822	29/30
		Static	49	−75	−3	348	28/30	−46	−76	−3	543	29/30

Face-selective areas were more robustly identified with dynamic stimuli in the pSTS and pFC. A contrast of dynamic scenes > dynamic objects produced large clusters encompassing the PPA, RSC, and large areas of the visual cortex in 26 participants. The object-selective LO also produced large clusters contiguous with early visual areas when defined using a contrast of moving objects > scrambled objects. The body-selective EBA was the most consistently identified ROI across participants, regardless of whether it was identified using dynamic or static stimuli. OFA = occipital face area.

static objects. We attempted to localize scene-selective voxels in the three commonly studied scene-selective ROIs. These were the parahippocampal place area (PPA), retrosplenial cortex (RSC), and the occipital place area (OPA). Results demonstrated that moving scenes greater than moving objects generated activations in the PPA and RSC that were contiguous not only with each other but also with large sections of visual cortex in 23 participants.

These large clusters yielded a large number of contiguous voxels, so that efficient functional localization needs to also rely on anatomical constraints or spatial constraints from existing group templates. The OPA was spatially distinct from this cluster in most participants (see Table 1). Static scenes greater than static objects demonstrated results more consistent with earlier fMRI studies of scene-selective ROIs. The PPA, RSC, and OPA were

successfully localized in the right hemisphere of most participants. These ROIs were less consistent in the left hemisphere, notably the RSC (see Table 1).

Body-selective ROIs were identified in two separate analyses, using a contrast of moving bodies greater than moving objects and a contrast of static bodies greater than static objects. We attempted to localize the two most studied body-selective ROIs, the extrastriate body area (EBA) and fusiform body area (FBA). Notably, the EBA was identified bilaterally across more participants using both moving and static contrasts than any other category-selective ROI (Table 1). Results further showed that the EBA was larger when defined using moving than static stimuli, but this was not the case with the FBA. This is consistent with prior evidence showing that lateral category-selective brain areas exhibit a greater response to moving stimuli more than static stimuli (Pitcher et al., 2019).

Finally, we defined the object-selective area LO using contrasts of moving objects > moving scrambled objects and a contrast of static objects > static scrambled objects. Results showed a similar pattern to the scene-selective ROIs, namely, that defining LO using moving stimuli produced large bilateral ROIs that were contiguous with early visual cortex (Table 1). By contrast, defining LO using static stimuli was more consistent with prior studies (Pitcher, Charles, Devlin, Walsh, & Duchaine, 2009; Malach et al., 1995).

DISCUSSION

In the present study, we used fMRI to measure the neural responses to dynamic and static stimuli from four visual categories (faces, bodies, scenes, and objects). Our aim was to establish the brain areas that selectively respond to moving stimuli of different categories across the three visual pathways. Results supported functional dissociations consistent with the recently proposed third visual pathway for social perception (Pitcher & Ungerleider, 2021). Specifically, we report the significant activity that projects into the STS for moving faces (Figure 4A) and moving bodies (Figure 4B) only. Moving objects (Figure 4C) and moving bodies (Figure 4C) both produced significant activity in the bilateral PPC, part of the dorsal visual pathway for visually guided action (Milner & Goodale, 1995). This pattern of results demonstrates that the motion of stimuli from high-level object categories is preferentially processed in the lateral and dorsal areas of the visual cortex, more than category-selective areas on the ventral brain surface. These results also suggest that there is a functional division between biological- and nonbiological stimuli across the dorsal and third visual pathways.

Preferential representations of dynamic face stimuli in the STS aligns with the role of the STS in social perception (Kilner, 2011; Allison et al., 2000; Perrett et al., 1992) and representing aspects of the face that can change rapidly such as expression, gaze, and mouth movements (Haxby et al., 2000). Accordingly, selectivity to facial motion

facilitates emotion perception in facial expressions and bodies (Atkinson, Dittrich, Gemmell, & Young, 2004; Kilts, Egan, Gideon, Ely, & Hoffman, 2003) and audiovisual integration of speech (Young, Fröhholz, & Schweinberger, 2020). The STS has also been implicated in biological motion perception, producing a greater response to motion stimuli depicting jumping, kicking, running, and throwing movements than control motion (Grossman et al., 2000). Such motion stimuli, as well as changeable aspects of the face, convey information that may provoke attributions of intentionality and personality of other individuals (Adolphs, 2002). Moreover, previous studies have shown that the STS exhibits high selectivity to social in contrast to nonsocial stimuli (Watson, Cardillo, Bromberger, & Chatterjee, 2014; Lahnakoski et al., 2012). In addition to finding that the STS seemed to be “people selective,” Watson and colleagues (2014) demonstrated the multisensory nature of the STS, proposing this region plays a vital role in combining socially relevant information across modalities. This discrimination between social and nonsocial stimuli is consistent with our proposal of a third pathway connecting V5 and STS, which is specialized for processing dynamic aspects of social perception (Pitcher & Ungerleider, 2021).

Our results for moving bodies are consistent with previous studies of body-selectivity in humans, which show that the EBA and the FBA respond more strongly to human bodies and body parts than faces, objects, scenes, and other stimuli (Peelen & Downing, 2005, 2007; Downing et al., 2001). The pSTS has also been implicated in the perception of biological motion through faces or bodies (Saygin, 2007; Puce & Perrett, 2003; Grossman et al., 2000). Previous studies have highlighted the dissociation in the response to dynamic and static presentation of bodies in lateral and ventral regions (Pitcher et al., 2019; Grosbras, Beaton, & Eickhoff, 2012) that we also report here. The location of these body-selective regions in distinct neuroanatomical pathways, with the FBA located on the ventral surface (Peelen & Downing, 2005; Schwarzlose, Baker, & Kanwisher, 2005) and the EBA (Downing et al., 2001) and pSTS (Grossman et al., 2000) on the lateral surface, is consistent with these regions having different response profiles to static and dynamic images of bodies. Interestingly, we also observed significant activity to body stimuli in the PPC, a brain area that is part of the dorsal processing stream for computing visually guided physical actions (Milner & Goodale, 1995). This pattern of results suggests that the multifaceted behavioral relevance of bodies triggers processing in both the dorsal and third visual pathways: Bodies are not only relevant for inferring social information about others (like faces), they also are critical for perceiving and evaluating visually guiding action (like objects). Our results are therefore consistent with the idea that the differential routing of categorical information across the dorsal and third pathways is not determined by movement per se but by the behavioral implications carried by the movement for a specific stimulus category.

Neuroimaging studies have identified multiple scene-selective brain regions in humans, including the PPA (Epstein & Kanwisher, 1998), the RSC (Maguire, Vargha-Khadem, & Mishkin, 2001), and the OPA (Dilks, Julian, Paunov, & Kanwisher, 2013). Previous studies have highlighted the OPA showing a greater response to dynamic than static scenes, whereas the PPA and RSC showed similar responses to dynamic and static scenes (Pitcher et al., 2019; Kamps, Lall, & Dilks, 2016; Korkmaz Hacialihafiz & Bartels, 2015). This is consistent with our result that moving scenes do not activate scene-selective areas in the ventral stream more strongly than static scenes. In addition, this selective response in the OPA to dynamic scenes is seemingly not as a result of low-level information processing or domain-general motion sensitivity (Kamps et al., 2016). The way these scene-selective regions can be dissociated based on motion sensitivity aligns with their possible roles in scene processing. Mirroring the role of the ventral pathway in recognition and the dorsal pathway in visually guided action (Milner & Goodale, 1995), these findings are consistent with the hypothesis of two distinct scene processing systems engaged in navigation and other aspects of scene processing such as scene categorization (Persichetti & Dilks, 2016; Dilks, Julian, Kubilius, Spelke, & Kanwisher, 2011). The anatomical position of the OPA within the dorsal pathway is compatible with its motion sensitivity and role in visually guided navigation (Kamps et al., 2016). Whereas the ventral/medial location of the PPA and RSC aligns with their demonstration of less sensitivity to motion and role in other aspects of navigation and scene recognition. Notably, our study revealed much more widespread scene-selective clusters when localization was performed with dynamic stimuli. These larger clusters may reflect the type of movement present in the scene stimuli: Rather than local movement of a foreground object (as present in the faces, body, and object stimuli), scene stimuli were characterized by more global movement patterns, where instead of the scene itself, the camera would move. Such global movement patterns may indeed be a relevant source of information in scene processing, where temporal dynamics in the information are often a consequence of the observer moving through the world. However, future studies need to systematically compare such global movements with scenes in which many local elements actively move (like trees and leaves move during a windy day) to delineate whether there are different consequences of these movement types on scene representation.

In addition to the group whole-brain analyses, we also performed ROI analyses for all four stimulus categories at the individual participant level (Table 1), allowing us to quantify whether moving stimuli can improve the quality of functional localization in the visual system. Prior fMRI studies have mostly used static images of stimuli from these categories to identify the relevant category-selective brain areas. However, a subset of areas are known to exhibit a greater neural response to moving more than static images from the preferred visual object category.

These include face-selective areas in the STS (Pitcher, Dilks, et al., 2011; Fox et al., 2009; LaBar et al., 2003; Puce et al., 1998), the scene-selective OPA in the transverse occipital sulcus (Kamps et al., 2016), and the body-selective EBA in the lateral occipital lobe (Pitcher et al., 2019). This spatial dissociation between moving and static stimuli was also observed in the ROI analyses performed for each individual participant (Table 1). Results were consistent with prior studies of these same areas (Sliwiska, Bearpark, Corkhill, McPhillips, & Pitcher, 2020; Pitcher et al., 2019; Pitcher, Dilks, et al., 2011). Our ROI results also demonstrate that moving stimuli successfully identify more ROIs across participants and larger ROIs than static stimuli for face, scene, and body areas, particularly in lateral brain areas. By contrast, moving stimuli did not lead to systematic shifts in ROI peak coordinates, showing that the activations indeed stem from the same cortical areas. The differential selectivity for motion is believed to relate to the different cognitive functions performed on stimuli within the relevant category, for example, facial identity or facial expression recognition (Haxby et al., 2000), or scene recognition or spatial navigation (Kamps et al., 2016). From a methodological perspective, the results shown in Table 1 also demonstrate that the use of moving stimuli for fMRI functional localizers will more robustly identify category-selective ROIs for all four stimulus categories across both hemispheres. The choice of stimuli in functional localizer experiments depends on multiple factors, including how these regions are probed in the subsequent experiments. Our study, however, provides a valuable benchmark for ROI-based fMRI studies in the future, where functional localization will be robustly achieved in a greater percentage of participants when moving stimuli are used instead of static stimuli.

This demonstration that moving stimuli produce larger and more robust ROIs also raises a methodological question about how to dissociate functionally selective brain areas from visual pathways. Or to phrase it differently, how do you simultaneously study the neural responses across a number of ROIs when these ROIs respond to a wide range of visual factors? One elegant solution to this issue has recently been demonstrated in a study that used fMRI to record neural activity while participants viewed short videos depicting a range of social interactions between two people (McMahon, Bonner, & Isik, 2023). The videos were annotated to identify visual features that should be selectively processed in brain areas at different levels of the visual hierarchy. These included low-level features (e.g., contrast and motion energy), mid-level features (e.g., physical distance between the actors and their direction of attention), and high-level features that support social understanding (e.g., the nature and valence of the interaction). Results showed that low-level features were preferentially processed in V5/MT, mid-level features were preferentially processed in the EBA and LO, whereas high-level features were preferentially processed in the STS. It is hoped that future studies using a similar

approach can identify the functions of specific brain areas (e.g., the motion-selective area or the STS) while also mapping the connectivity and broader functionality between these brain areas. We believe that visual pathway models create the broader framework in which these results can be understood and interpreted.

Acknowledgments

Thanks to Nancy Kanwisher for providing experimental stimuli; and thanks to Shruti Japee for help with data analysis.

Corresponding author: David Pitcher, Department of Psychology, University of York, Heslington, York, YO10 5DD, UK, or via e-mail: david.pitcher@york.ac.uk.

Data Availability Statement

Data for all participants are available on request. The whole-brain group activation maps used to generate the images in Figures 3, 4, and 5 are available at <https://osf.io/3w4ps/>.

Author Contributions

Emel Küçük: Methodology; Writing—Original draft. Matthew Foxwell: Investigation; Writing—Original draft. Daniel Kaiser: Data curation; Formal analysis; Methodology; Writing—Original draft. David Pitcher: Conceptualization; Data curation; Formal analysis; Methodology; Writing—Original draft.

Funding Information

This work was funded by a grant from the Biotechnology and Biological Sciences Research Council (<https://dx.doi.org/10.13039/501100000268>), grant number: BB/P006981/1 (awarded to D. P.). D. K. is supported by the German Research Foundation (DFG; <https://dx.doi.org/10.13039/501100001659>), grant numbers: SFB/TRR135, Project Number 222641018; KA4683/5-1, Project Number 518483074, by “The Adaptive Mind”, funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art, and an ERC (<https://dx.doi.org/10.13039/501100000781>) Starting Grant (PEP, ERC-2022-STG 101076057). Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period

were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article’s gender citation balance.

REFERENCES

- Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, 1, 21–62. <https://doi.org/10.1177/1534582302001001003>, PubMed: 17715585
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, 4, 267–278. [https://doi.org/10.1016/s1364-6613\(00\)01501-1](https://doi.org/10.1016/s1364-6613(00)01501-1), PubMed: 10859571
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33, 717–746. <https://doi.org/10.1068/p5096>, PubMed: 15330366
- Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). fMRI responses to video and point-light displays of moving humans and manipulable objects. *Journal of Cognitive Neuroscience*, 15, 991–1001. <https://doi.org/10.1162/089892903770007380>, PubMed: 14614810
- Beauchamp, M. S., Nath, A. R., & Pasalar, S. (2010). fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. *Journal of Neuroscience*, 30, 2414–2417. <https://doi.org/10.1523/JNEUROSCI.4865-09.2010>, PubMed: 20164324
- Boussaoud, D., Ungerleider, L. G., & Desimone, R. (1990). Pathways for motion analysis: Cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *Journal of Comparative Neurology*, 296, 462–495. <https://doi.org/10.1002/cne.902960311>, PubMed: 2358548
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77, 305–327. <https://doi.org/10.1111/j.2044-8295.1986.tb02199.x>, PubMed: 3756376
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, 6, 641–651. <https://doi.org/10.1038/nrn1724>, PubMed: 16062171
- Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., et al. (1997). Activation of auditory cortex during silent lipreading. *Science*, 276, 593–596. <https://doi.org/10.1126/science.276.5312.593>, PubMed: 9110978
- Campbell, R., Heywood, C. A., Cowey, A., Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia*, 28, 1123–1142. [https://doi.org/10.1016/0028-3932\(90\)90050-x](https://doi.org/10.1016/0028-3932(90)90050-x), PubMed: 2290489
- de Haan, E. H. F., & Cowey, A. (2011). On the usefulness of ‘what’ and ‘where’ pathways in vision. *Trends in Cognitive Sciences*, 15, 460–466. <https://doi.org/10.1016/j.tics.2011.08.005>, PubMed: 21906989
- Desimone, R., & Ungerleider, L. G. (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. *Journal of Comparative Neurology*, 248,

- 164–189. <https://doi.org/10.1002/cne.902480203>, PubMed: 3722457
- Dilks, D. D., Julian, J. B., Kubilius, J., Spelke, E. S., & Kanwisher, N. (2011). Mirror-image sensitivity and invariance in object and scene processing pathways. *Journal of Neuroscience*, *31*, 11305–11312. <https://doi.org/10.1523/JNEUROSCI.1935-11.2011>, PubMed: 21813690
- Dilks, D. D., Julian, J. B., Paunov, A. M., & Kanwisher, N. (2013). The occipital place area is causally and selectively involved in scene perception. *Journal of Neuroscience*, *33*, 1331–1336. <https://doi.org/10.1523/JNEUROSCI.4081-12.2013>, PubMed: 23345209
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*, 2470–2473. <https://doi.org/10.1126/science.1063414>, PubMed: 11577239
- Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, *1*, 393–416. <https://doi.org/10.1146/annurev-vision-082114-035518>, PubMed: 28532371
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601. <https://doi.org/10.1038/33402>, PubMed: 9560155
- Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, *12*, 388–396. <https://doi.org/10.1016/j.tics.2008.07.004>, PubMed: 18760955
- Fodor, J. A. (1983). *The modularity of mind*. MIT Press. <https://doi.org/10.7551/mitpress/4737.001.0001>
- Fox, C. J., Iaria, G., & Barton, J. J. S. (2009). Defining the face processing network: Optimization of the functional localizer in fMRI. *Human Brain Mapping*, *30*, 1637–1651. <https://doi.org/10.1002/hbm.20630>, PubMed: 18661501
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649–677. <https://doi.org/10.1146/annurev.neuro.27.070203.144220>, PubMed: 15217346
- Grosbras, M.-H., Beaton, S., & Eickhoff, S. B. (2012). Brain regions involved in human movement perception: A quantitative voxel-based meta-analysis. *Human Brain Mapping*, *33*, 431–454. <https://doi.org/10.1002/hbm.21222>, PubMed: 21391275
- Grossman, E. D., & Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron*, *35*, 1167–1175. [https://doi.org/10.1016/S0896-6273\(02\)00897-8](https://doi.org/10.1016/S0896-6273(02)00897-8), PubMed: 12354405
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., et al. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, *12*, 711–720. <https://doi.org/10.1162/089892900562417>, PubMed: 11054914
- Handwerker, D. A., Ianni, G., Gutierrez, B., Roopchansingh, V., Gonzalez-Castillo, J., Chen, G., et al. (2020). Theta-burst TMS to the posterior superior temporal sulcus decreases resting-state fMRI connectivity across the face processing network. *Network Neuroscience*, *4*, 746–760. https://doi.org/10.1162/netn_a_00145, PubMed: 32885124
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233. [https://doi.org/10.1016/s1364-6613\(00\)01482-0](https://doi.org/10.1016/s1364-6613(00)01482-0), PubMed: 10827445
- Hein, G., & Knight, R. T. (2008). Superior temporal sulcus—It's my area: Or is it? *Journal of Cognitive Neuroscience*, *20*, 2125–2136. <https://doi.org/10.1162/jocn.2008.20148>, PubMed: 18457502
- Kamps, F. S., Lall, V., & Dilks, D. D. (2016). The occipital place area represents first-person perspective motion information through scenes. *Cortex*, *83*, 17–26. <https://doi.org/10.1016/j.cortex.2016.06.022>, PubMed: 27474914
- Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences, U.S.A.*, *107*, 11163–11170. <https://doi.org/10.1073/pnas.1005062107>, PubMed: 20484679
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311. <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>, PubMed: 9151747
- Kilner, J. M. (2011). More than one pathway to action understanding. *Trends in Cognitive Sciences*, *15*, 352–357. <https://doi.org/10.1016/j.tics.2011.06.005>, PubMed: 21775191
- Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *Neuroimage*, *18*, 156–168. <https://doi.org/10.1006/nimg.2002.1323>, PubMed: 12507452
- Korkmaz Hacialihafiz, D., & Bartels, A. (2015). Motion responses in scene-selective regions. *Neuroimage*, *118*, 438–444. <https://doi.org/10.1016/j.neuroimage.2015.06.031>, PubMed: 26091852
- Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience*, *12*, 217–230. <https://doi.org/10.1038/nrn3008>, PubMed: 21415848
- Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, *17*, 26–49. <https://doi.org/10.1016/j.tics.2012.10.011>, PubMed: 23265839
- LaBar, K. S., Crupain, M. J., Voyvodic, J. T., & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex*, *13*, 1023–1033. <https://doi.org/10.1093/cercor/13.10.1023>, PubMed: 12967919
- Lahnakoski, J. M., Glerean, E., Salmi, J., Jääskeläinen, I. P., Sams, M., Hari, R., et al. (2012). Naturalistic fMRI mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. *Frontiers in Human Neuroscience*, *6*, 233. <https://doi.org/10.3389/fnhum.2012.00233>, PubMed: 22905026
- Maguire, E. A., Vargha-Khadem, F., & Mishkin, M. (2001). The effects of bilateral hippocampal damage on fMRI regional activations and interactions during memory retrieval. *Brain*, *124*, 1156–1170. <https://doi.org/10.1093/brain/124.6.1156>, PubMed: 11353732
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *92*, 8135–8139. <https://doi.org/10.1073/pnas.92.18.8135>, PubMed: 7667258
- McMahon, E., Bonner, M. F., & Isik, L. (2023). Hierarchical organization of social action features along the lateral visual pathway. *Current Biology*, *33*, 5035–5047. <https://doi.org/10.1016/j.cub.2023.10.015>, PubMed: 37918399
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford University Press.
- Nikel, L., Sliwinska, M. W., Kucuk, E., Ungerleider, L. G., & Pitcher, D. (2022). Measuring the response to visually presented faces in the human lateral prefrontal cortex. *Cerebral Cortex Communications*, *3*, tgac036. <https://doi.org/10.1093/texcom/tgac036>, PubMed: 36159205

- O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences*, 6, 261–266. [https://doi.org/10.1016/S1364-6613\(02\)01908-3](https://doi.org/10.1016/S1364-6613(02)01908-3), PubMed: 12039608
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, 93, 603–608. <https://doi.org/10.1152/jn.00513.2004>, PubMed: 15295012
- Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience*, 8, 636–648. <https://doi.org/10.1038/nrn2195>, PubMed: 17643089
- Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 335, 23–30. <https://doi.org/10.1098/rstb.1992.0003>, PubMed: 1348133
- Persichetti, A. S., & Dilks, D. D. (2016). Perceived egocentric distance sensitivity and invariance across scene-selective cortex. *Cortex*, 77, 155–163. <https://doi.org/10.1016/j.cortex.2016.02.006>, PubMed: 26963085
- Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., et al. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature*, 389, 495–498. <https://doi.org/10.1038/39051>, PubMed: 9333238
- Pilz, K. S., Vuong, Q. C., Bülthoff, H. H., & Thornton, I. M. (2011). Walk this way: Approaching bodies can influence the processing of faces. *Cognition*, 118, 17–31. <https://doi.org/10.1016/j.cognition.2010.09.004>, PubMed: 21047624
- Pitcher, D. (2021). Characterizing the third visual pathway for social perception. *Trends in Cognitive Sciences*, 25, 550–551. <https://doi.org/10.1016/j.tics.2021.04.008>, PubMed: 34024729
- Pitcher, D., Charles, L., Devlin, J. T., Walsh, V., & Duchaine, B. (2009). Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Current Biology*, 19, 319–324. <https://doi.org/10.1016/j.cub.2009.01.007>, PubMed: 19200723
- Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *Neuroimage*, 56, 2356–2363. <https://doi.org/10.1016/j.neuroimage.2011.03.067>, PubMed: 21473921
- Pitcher, D., Duchaine, B., & Walsh, V. (2014). Combined TMS and fMRI reveal dissociable cortical pathways for dynamic and static face perception. *Current Biology*, 24, 2066–2070. <https://doi.org/10.1016/j.cub.2014.07.060>, PubMed: 25131678
- Pitcher, D., Ianni, G., & Ungerleider, L. G. (2019). A functional dissociation of face-, body- and scene-selective brain areas based on their response to moving and static stimuli. *Scientific Reports*, 9, 8242. <https://doi.org/10.1038/s41598-019-44663-9>, PubMed: 31160680
- Pitcher, D., Pilkington, A., Rauth, L., Baker, C., Kravitz, D. J., & Ungerleider, L. G. (2020). The human posterior superior temporal sulcus samples visual space differently from other face-selective regions. *Cerebral Cortex*, 30, 778–785. <https://doi.org/10.1093/cercor/bhz125>, PubMed: 31264693
- Pitcher, D., Sliwiska, M. W., & Kaiser, D. (2023). TMS disruption of the lateral prefrontal cortex increases neural activity in the default mode network when naming facial expressions. *Social Cognitive and Affective Neuroscience*, 18, nsad072. <https://doi.org/10.1093/scan/nsad072>, PubMed: 38048419
- Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a third visual pathway specialized for social perception. *Trends in Cognitive Sciences*, 25, 100–110. <https://doi.org/10.1016/j.tics.2020.11.006>, PubMed: 33334693
- Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental Brain Research*, 209, 481–493. <https://doi.org/10.1007/s00221-011-2579-1>, PubMed: 21318346
- Pourtois, G., Sander, D., Andres, M., Grandjean, D., Reveret, L., Olivier, E., et al. (2004). Dissociable roles of the human somatosensory and superior temporal cortices for processing social face signals. *European Journal of Neuroscience*, 20, 3507–3515. <https://doi.org/10.1111/j.1460-9568.2004.03794.x>, PubMed: 15610183
- Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, 18, 2188–2199. <https://doi.org/10.1523/JNEUROSCI.18-06-02188.1998>, PubMed: 9482803
- Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 358, 435–445. <https://doi.org/10.1098/rstb.2002.1221>, PubMed: 12689371
- Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain*, 130, 2452–2461. <https://doi.org/10.1093/brain/awm162>, PubMed: 17660183
- Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research*, 194, 465–475. <https://doi.org/10.1007/s00221-009-1721-9>, PubMed: 19205678
- Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *Journal of Neuroscience*, 25, 11055–11059. <https://doi.org/10.1523/JNEUROSCI.2621-05.2005>, PubMed: 16306418
- Sliwiska, M. W., Bearpark, C., Corkhill, J., McPhillips, A., & Pitcher, D. (2020). Dissociable pathways for moving and static face perception begin in early visual cortex: Evidence from an acquired prosopagnosic. *Cortex*, 130, 327–339. <https://doi.org/10.1016/j.cortex.2020.03.033>, PubMed: 32736196
- Sliwiska, M. W., Elson, R., & Pitcher, D. (2020). Dual-site TMS demonstrates causal functional connectivity between the left and right posterior temporal sulci during facial expression recognition. *Brain Stimulation*, 13, 1008–1013. <https://doi.org/10.1016/j.brs.2020.04.011>, PubMed: 32335230
- Sliwiska, M. W., & Pitcher, D. (2018). TMS demonstrates that both right and left superior temporal sulci are important for facial expression recognition. *Neuroimage*, 183, 394–400. <https://doi.org/10.1016/j.neuroimage.2018.08.025>, PubMed: 30130641
- Sliwiska, M. W., Searle, L. R., Earl, M., O'Gorman, D., Pollicina, G., Burton, A. M., et al. (2022). Face learning via brief real-world social interactions induces changes in face-selective brain areas and hippocampus. *Perception*, 51, 521–538. <https://doi.org/10.1177/03010066221098728>, PubMed: 35542977
- Ungerleider, L. G., & Desimone, R. (1986). Cortical connections of visual area MT in the macaque. *Journal of Comparative Neurology*, 248, 190–222. <https://doi.org/10.1002/cne.902480204>, PubMed: 3722458
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In *Analysis of Visual Behavior* (pp. 549–586). Cambridge, MA: MIT Press.
- Walsh, V., & Butler, S. R. (1996). Different ways of looking at seeing. *Behavioural Brain Research*, 76, 1–3. [https://doi.org/10.1016/0166-4328\(96\)00189-1](https://doi.org/10.1016/0166-4328(96)00189-1), PubMed: 8734039

- Watson, C. E., Cardillo, E. R., Bromberger, B., & Chatterjee, A. (2014). The specificity of action knowledge in sensory and motor systems. *Frontiers in Psychology*, 5, 494. <https://doi.org/10.3389/fpsyg.2014.00494>, PubMed: 24904506
- Watson, J. D., Myers, R., Frackowiak, R. S., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., et al. (1993). Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, 3, 79–94. <https://doi.org/10.1093/cercor/3.2.79>, PubMed: 8490322
- Young, A. W., Fröhholz, S., & Schweinberger, S. R. (2020). Face and voice perception: Understanding commonalities and differences. *Trends in Cognitive Sciences*, 24, 398–410. <https://doi.org/10.1016/j.tics.2020.02.001>, PubMed: 32298625