

Opinion

Two ‘what’ pathways for action and object recognition

Moritz F. Wurm ^{1,*} and Alfonso Caramazza^{1,2}

The ventral visual stream is conceived as a pathway for object recognition. However, we also recognize the actions an object can be involved in. Here, we show that action recognition critically depends on a pathway in lateral occipitotemporal cortex, partially overlapping and topographically aligned with object representations that are precursors for action recognition. By contrast, object features that are more relevant for object recognition, such as color and texture, are typically found in ventral occipitotemporal cortex. We argue that occipitotemporal cortex contains similarly organized lateral and ventral ‘what’ pathways for action and object recognition, respectively. This account explains a number of observed phenomena, such as the duplication of object domains and the specific representational profiles in lateral and ventral cortex.

Is there a ‘what’ stream for action recognition?

The ventral stream, specifically **ventral occipitotemporal cortex (VOTC)** (see [Glossary](#)), has been conceptualized as a pathway for the **recognition** of objects [1,2] (e.g., *scissors*, *my grandmother*, or *the Eiffel Tower*). Broadly, the ventral stream is organized along two orthogonal axes: a posterior–anterior gradient mediating the transformation from perceptual to more general, conceptual representations [3] and a medial–lateral specialization for different, evolutionarily salient knowledge domains such as animate and inanimate objects [4,5] and more specific object categories like faces, body parts, tools, and places [5–7]. This categorical specialization is thought to be driven by the connectivity to different downstream regions, for example, networks for social cognition, object manipulation, and navigation [8,9].

However, we do not only recognize objects (what is it?) but also the **actions** an object can be involved in (what is it *doing*?; e.g., *cutting*, *running*, or *greeting*). Recognizing actions (and perhaps also inanimate and unintentional events, e.g., *colliding* or *slipping*) plays a key role in guiding our behavior, as it provides the basis for inferring others’ plans and desires, reacting appropriately, and social learning. Hence, it seems plausible that evolutionary pressures led to the development of neural circuits specialized for action recognition. Here, we use a functional approach to address the following questions: Which parts of the brain form a what pathway or pathways for action recognition? Do similar organizational principles apply to actions as has been observed for objects? Since action recognition usually builds on the recognition of objects, in that actions are constrained by the properties of objects as agents, instruments, or recipients of actions, how do the neural representations of objects (e.g., *body parts*, *tools*, or *fruits*) and actions (e.g., *running*, *cutting*, and *eating*) relate to each other? To address these questions, we focus on action aspects that capture what an action is for; typically the effect of an action (e.g., *opening* as a means to bring something into an open state), and leave aside other kinds of action knowledge, such as knowledge related to action production (how to perform a certain action, e.g., with the hand or foot) or the hidden motivations and reasons of actions (why an action is performed, e.g., *drinking to quench thirst*), which may rely on complementary neural systems.

Highlights

Object recognition draws on a ventral what pathway. Is there also a what pathway for action recognition? Recent neuroimaging findings suggest that critical stages of action recognition draw on a pathway in lateral occipitotemporal cortex.

This lateral action pathway is organized along two main dimensions: a posterior–anterior gradient from perceptual action precursors to conceptual action representations and a dorsal–ventral distinction that segregates animate and inanimate action aspects. This organization mirrors the domain-specific organization of the ventral object pathway.

We propose that occipitotemporal cortex subdivides into lateral and ventral what pathways for action and object recognition. This view explains several observed phenomena of functional organization and thereby extends and concretizes previous conceptions of pathways in visual cortex.

¹Center for Mind/Brain Sciences – CIMeC, University of Trento, Corso Bettini 31, 38068 Rovereto, Italy

²Department of Psychology, Harvard University, 33 Kirkland St, Cambridge, MA 02138, USA

*Correspondence: moritz.wurm@unitn.it (M.F. Wurm).

While action recognition draws on a network of occipitotemporal, parietal, and frontal brain regions, recent proposals emphasized the role of the **lateral occipitotemporal cortex (LOTc)** in representing various action-related aspects such as tools and body parts and their movements, verbs, and higher-level semantic action knowledge [10–12]. The LOTc is located in a pathway that appears structurally separate from the traditional dorsal and ventral pathways [12–15]. In the following, we build on these foundations and integrate the different functional roles proposed for this pathway with recent findings pointing toward an action recognition hierarchy in the LOTc that mirrors the domain-specific organization of the ventral object pathway. We argue that the original conception of the ventral stream as a what pathway can be extended to the LOTc. According to this view, lateral and ventral parts of the OTc subserve complementary recognition functions, with subordinate specializations for action and object recognition, respectively. This interpretation explains several observed phenomena, such as the dissociation of response preferences for object properties that are disproportionately relevant for action or object recognition (e.g., motion and color, respectively) and the duplication of response preferences for certain object categories that are important for both action and object recognition (e.g., body parts, faces, and inanimate objects).

Neural representations for action recognition in the LOTc

Action recognition is associated with activation of a network consisting of the LOTc and lateral frontal and parietal brain regions [10,16,17]. Since recognizing an action (e.g., *cutting an apple*) usually includes the visual analysis of objects (*apple, knife, and hands*), a straightforward interpretation is that the LOTc processes perceptual action details like body parts and objects and their movements whereas higher-level *conceptual* action aspects are encoded more anteriorly and dorsally in frontoparietal cortex. However, findings taken to support this view also allow alternative interpretations [16,18], and other findings – in particular those based on the performance of brain damaged individuals [19,20] – provide counterevidence against this view (but see [21]). Instead of reviewing the arguments resulting from this large body of work, we focus in this section on more recent neuroimaging studies which aimed at characterizing the representational profiles of occipitotemporal and frontoparietal regions with specific regard to conceptual action representation, that is, the ‘cardinal stage’ of action recognition that captures the core ‘what’ meaning of an action. We then use this stage as an anchor point from which we move posteriorly to characterize more basic, perceptual levels of action representation.

Conceptual action representations

Which parts of the brain represent actions at a conceptual level? We define three criteria that seem necessary (but not sufficient in isolation): (i) an action concept (e.g., *opening*) should generalize across different exemplars of the action (e.g., the opening of a *bottle* or of a *box*); (ii) an action concept should generalize across different input modalities or stimulus formats, for example, it should respond to both the observation of an action and to a verbal description of that action; and (iii) the similarity of conceptual action representations should follow specific semantic principles (see [Box 1](#) for further details). Recent advances in neuroimaging, specifically **multivariate pattern analysis (MVPA)** and representational similarity analysis (RSA), allow identifying brain regions that meet these criteria. fMRI studies using these approaches revealed exemplar-general action representations in the anterior and posterior LOTc and inferior parietal lobe (IPL). Action representations in these regions generalize across perceptual variants of these actions differing in terms of the objects involved, for example, bottles, boxes, doors, and trash bins [22–25], and the body parts [22] and their movements [25] needed to manipulate these different objects. Frontal brain regions, such as premotor cortex, are mostly found for the discrimination of actions that differ not only conceptually but also in terms of involved body parts [26] or body movements [24], suggesting that frontal representations encode effector- and movement-related rather than

Glossary

Action: purposeful behavior that has a certain impact on the world (e.g., on objects, others, the self) and thereby serves to achieve a certain goal. We broadly distinguish physical actions that are carried out motorically (e.g., kicking, cutting, and greeting) from nonphysical, typically mental actions (e.g., thinking and planning). We distinguish what an action does (the effect of an action) from how an action does it (or the means of how an action is carried out) and why an action is carried out (the underlying, hidden motivation or intention of an action).

Goal: the outcome that is aimed to be achieved by an action, typically a certain state of an object, another person, or the self. Goals refer to mental states of the acting person and should therefore not be confused with actions. Goals usually match the actually performed action, but not necessarily (e.g., she has the goal to bend the stick, but accidentally breaks it).

Lateral occipitotemporal cortex (LOTc): coarsely defined cortical region centered around the lateral occipital sulcus, which separates the anterior LOTc (i.e., lateral posterior temporal cortex) from posterior LOTc (i.e., lateral occipital cortex). We conceive the LOTc as part of the lateral visual pathway (or stream), which emerges from the primary visual cortex, continues through the LOTc, and eventually reaches the lateral temporal cortex.

Multivariate pattern analysis (MVPA): a method used in neuroimaging to discriminate neural activation patterns associated with different representations, for example, of recognized actions.

Object: we use the term object in a context-dependent manner. In a general context, we refer to objects as both animate or inanimate entities in the world (e.g., person, dog, apple, knife, and house). With this definition, we differentiate, for example, object from action recognition. In the context of objects as agents, instruments, or recipients of actions, we refer to objects exclusively as inanimate entities. Thus, actions can be directed at, or involve the interaction with, persons (or other animate entities with mental states) or (inanimate) objects.

Recognition: the process that enables the match of a percept with a memory content, which broadly consist of the perceptual analysis of a stimulus (e.g.,

Box 1. Localizing conceptual action representations in the brain

Neural representations of action concepts should fulfil three criteria. (i) An action concept (e.g., *opening*) should generalize across different concrete instantiations of the action type (Figure 1A, left). Note that this criterion is stricter than viewpoint or movement invariance; for example, when observing the same action from different viewpoints or when the same action outcome is achieved with different body movements. Representations that generalize across the latter dimensions could also capture higher-level perceptual features such as viewpoint-invariant motion trajectories or shape. Exemplar-general action representations can be identified by training a classifier to distinguish neural activation patterns associated with the observation of different actions (e.g., of *opening* and *closing bottles*) and testing the classifier on its ability to distinguish the same kinds of actions on different objects (e.g., *boxes*). (ii) An action concept should generalize across different input modalities; for example, observation or language (Figure 1A, middle). Modality-general action representations can be identified by training a classifier to discriminate activation patterns associated with observed actions and testing the classifier on its accuracy to discriminate activation patterns associated with corresponding action sentences. (iii) The representational similarity of different action concepts should reflect their semantic similarity rather than similarity along concept-irrelevant dimensions (Figure 1A, right). Among different semantic dimensions, we here refer to ‘what’ aspects that are critical for defining an action’s meaning. For example, the core ‘what’ meaning of *hitting* is the (forceful) contact with another object rather than the fact that it can be carried out with the foot or with the hand, which reflect narrower meanings related to ‘how’ aspects. Using RSA, the similarities between activation patterns associated with different actions can be compared with models of similarity that capture, for example, whether actions are directed toward persons or involve an interaction with manipulable objects. fMRI studies that were specifically designed to localize neural representations that meet these criteria reveal converging evidence for conceptual action representation in left anterior LOTC (Figure 1B).

object shapes and movement patterns) and the access of more general, conceptual knowledge. This memory content can also be accessed without perception, for example, through language. At the neural level, this match is reflected by the activation of a specific neural population – the activated neural population therefore represents the content (at a certain level of generality), which implies discrimination from other contents. Operationally, discrimination can be tested using MVPA.

Ventral occipitotemporal cortex (VOTC): cortical region defined by the following lateral, medial, posterior, and anterior boundaries: occipitotemporal sulcus, parahippocampal gyrus, posterior transverse collateral sulcus, and anterior end of the fusiform sulcus. The VOTC is conceived as part of the ventral visual pathway (or stream), which is thought to project from primary visual cortex along VOTC to the temporal pole.

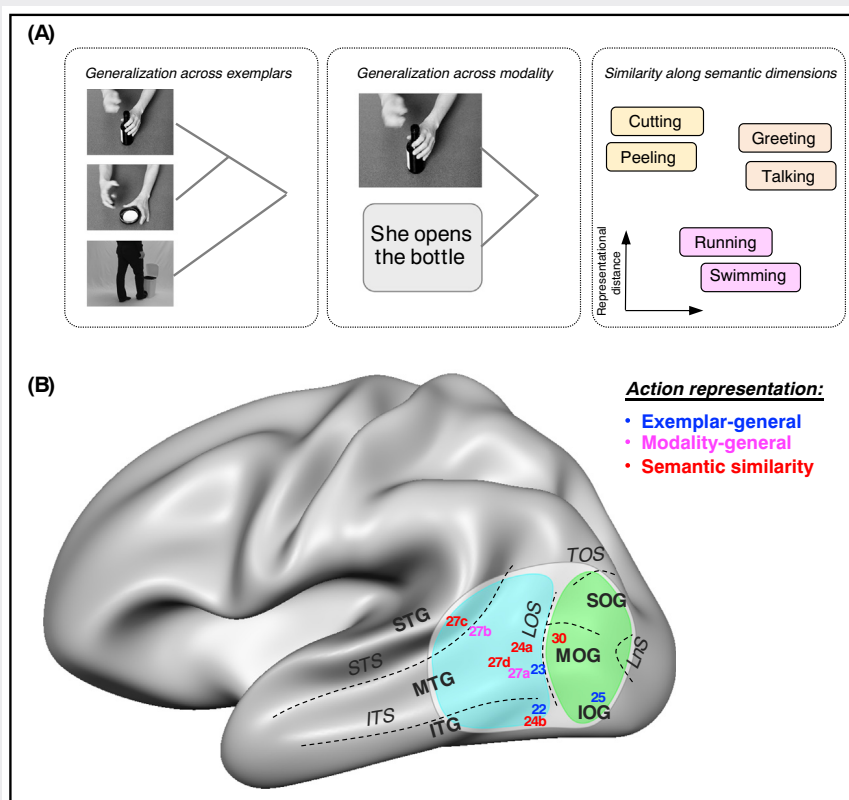


Figure 1. Three criteria for conceptual action representation. (A) Schematic overview of the criteria. (B) Results of multivariate pattern analysis (MVPA) studies that tested these criteria (peak coordinates). Numbers correspond to references listed in this article; letters refer to separate analyses/experiments within the same study. The white area indicates the coarse boundaries of the lateral occipitotemporal cortex (LOTC); light blue and green subregions indicate anterior and posterior LOTC, respectively. Abbreviations: IOG, inferior occipital gyrus; ITG/ITS, inferior temporal gyrus/sulcus; Lns, Lunate sulcus; LOS, lateral occipital sulcus; MOG, middle occipital gyrus; MTG/MTS, middle temporal gyrus/sulcus; SOG, superior occipital gyrus; STG/STS, superior temporal gyrus/sulcus; TOS, transverse occipital sulcus.

conceptual action information. Modality-general action representations, which are accessible both via observation and written descriptions, were found using fMRI-based crossdecoding in the left anterior LOTC [27]. In the same study, RSA revealed that the neural similarity of modality-general action representations in the anterior LOTC can be predicted by models that capture various semantic aspects: whether actions are directed toward persons or not, directed toward inanimate objects or not, or similar in terms of more complex semantic relations. This finding is in line with other studies investigating the representational organization of observed actions: Person and object directedness (also labeled sociality and transitivity, respectively) captures the similarity of activation patterns associated with observed actions in the LOTC [24,28]. The latter studies also found that frontoparietal and VOTC regions are sensitive to the person and object directedness of actions. However, only the left LOTC is found to be sensitive to semantic action dimensions when perceptual factors that typically covary with person and object directedness, such as reaching/grasping movements and presence of persons and objects, are explicitly modeled and thus controlled for [29,30]. These findings suggest that conceptual action representations are primarily represented in the left LOTC, whereas frontoparietal regions and the VOTC encode stimulus-specific action-related aspects. It is nonetheless possible that other regions meet conceptual criteria but are not detected with current neuroimaging methods. For example, single-cell recordings in humans have revealed that, in parietal cortex, action representations can be found that generalize across action videos and verbs [31]. However, the neural response profile found in this study were best explained by a model that captured stimulus-specific aspects, pointing toward a role of parietal cortex that is different from conceptual representation (see also [32]) (Box 2 for possible roles).

Taken together, these studies demonstrate that the left anterior LOTC meets all three criteria for conceptual action representation. This interpretation is in line with evidence from transcranial magnetic stimulation (TMS) studies [33], single-cell recordings in monkeys [34], and neuropsychological evidence [19], although further research is needed to directly compare the causal roles of the frontoparietal cortex and LOTC in accessing conceptual action representations. Another open question concerns the selectivity of the LOTC for the representation of action concepts; it could be that other kinds of information are also represented in this region. The

Box 2. The role of frontoparietal cortex in action recognition

The frontoparietal cortex, specifically premotor cortex and IPL, are typically activated during action recognition. The neuroimaging evidence presented here argues against the view that these brain regions represent actions at conceptual levels that captures the core ‘what’ meaning of actions. In addition, there is accumulating evidence against the view that sensorimotor representations of actions in frontoparietal cortex are strictly necessary for action recognition [18,20,111–114]. What then could be the role of frontal and parietal regions in action recognition and representation? Various studies suggest that frontoparietal regions, in particular anterior IPL, are necessary for knowing how a tool is grasped and manipulated for proper use [104,115]. In line with this view, the frontoparietal cortex appears mostly sensitive to the kinematics of actions [111]. Besides these ‘how’ aspects of action knowledge, there are also ‘why’ aspects, which capture the hidden motives underlying an action, such as desires, **goals**, and intentions. These inferred aspects of actions might be represented in frontoparietal regions, specifically in posterior IPL [116] or regions of the mentalizing network such as the temporoparietal junction and anterior STS [117,118]. Additionally, frontoparietal regions are likely involved in action recognition by providing supportive functions related to working memory [119] and visuospatial and attentional processes that are disproportionately relevant for action recognition relative to object recognition [16]. According to a recent proposal, IPL might also play a role in representing the ‘social affordances’ of actions, by integrating observed body movements and the effects of an action (originating from the LOTC) with potential motor plans to support the selection of behavioral reactions in frontal cortex [17]. Finally, frontoparietal regions could be involved in processes that follow action recognition such as activating associated motor representations in the service of learning and imitation [113], and anticipating future actions and events based on the recognized action [120]. Some of these functions also relate to mechanical reasoning, which may be particularly important for inferring the meaning of unfamiliar actions or functions of novel tools [11,104]. Related to this point, understanding how different entities in an observed action scene causally impact with each other might be represented at a perceptually invariant level and independent of object identity. This kind of representation might explain the finding of exemplar-general, but not modality-general action decoding in parietal cortex [22–25,27].

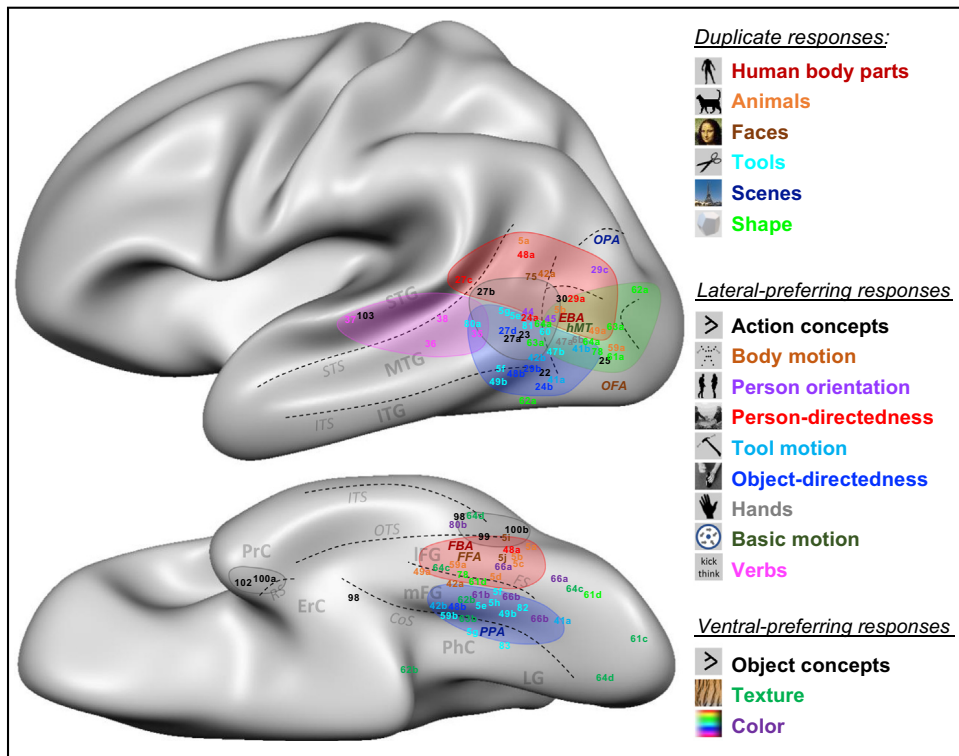
anterior LOTC comprises a region around the middle temporal gyrus (MTG) and superior temporal sulcus (STS) that responds to verbs (relative to matched nouns) [35], including verbs referring to nonphysical actions (e.g., *think*) [36] and states (e.g., *exist*) [37], pointing toward a more general role of the lateral temporal cortex in representing predicate structures that capture what happens and what is [38]. The anatomical location of the anterior LOTC between the visual cortex and language-related MTG/STS appears ideal for the representation of such structures.

Perceptual action precursors

At its posterior end, the 'conceptual action' LOTC borders on regions that are disproportionately activated by entities like body parts [39], tools, and other manipulable objects [4–6,40], and typical movements associated with them [41–43]. The posterior LOTC is also sensitive to visuospatial interobject relations, for example, how persons are oriented toward each other [29,44,45]. Likewise, person–object pairs can be decoded over and above collapsed responses to persons and objects in isolation, pointing toward the integration of person and object information into a single functional unit in the LOTC [46]. Moreover, an anterior–ventral subregion of the extrastriate body area (EBA) shows representational overlap between hands and tools [6,40,47] suggesting a representation of hand–object interaction that may function as a precursor for the recognition of object-directed actions. Finally, representations in the posterior LOTC are sensitive to the directedness of an action toward different classes of targets such as persons or inanimate objects (e.g., moving a cup towards a person or towards a saucer), over and above movements *per se* as well as the presence of these different entities in the action scene [29]. These more basic types of representation in the posterior LOTC are necessary for action recognition and thus can be considered perceptual precursors of higher-level, conceptual action representations in the anterior LOTC. The functional distinction between the posterior and anterior LOTC suggests a hierarchy of processing, from more basic perceptual to more integrated conceptual action representation.

Topographical alignment of action and object representations in the LOTC

For each level of representation – single entities, their movements, action directedness, and conceptual action aspects – categorical preferences are observed in dorsal and ventral subregions of the LOTC: sensitivity to animate entities, body parts, their canonical movements, orientation of persons relative to each other, and actions directed toward persons is preferentially found in dorsal rather than ventral subregions of the LOTC, whereas sensitivity to inanimate entities like tools, action-specific tool motion, and actions directed towards or involving manipulable objects is found in ventral rather than dorsal subregions of the LOTC [4,5,24,27,29,41,42,48–51] (Figure 1). The domain-dependent alignment of action and object representations can be explained by the specific roles that objects play in different kinds of actions and their importance for action identification: Recognizing social, person-directed actions like *greeting* relies on information such as body posture and movements, person orientation, as well as other socially relevant cues like facial expressions. By contrast, recognizing object-directed actions like *cutting* relies more on the recognition of inanimate objects and their configurations and state changes. Likewise, manipulable objects in isolation (as well as their properties like shape, material, weight, etc.) can trigger which actions can be done with them, suggesting a tight association between objects and the actions they afford. Based on the notion that functional organization in the brain is determined by the connectivity between associated regions [8] it follows plausibly that associated action and object representations are encoded along a common pathway. The topographical alignment thus suggests that aspects of action and object knowledge depend on each other in anterior and posterior LOTC, consistent with the view that object representations in LOTC capture the roles that objects play as agents, instruments, and recipients of actions and events.



Trends in Cognitive Sciences

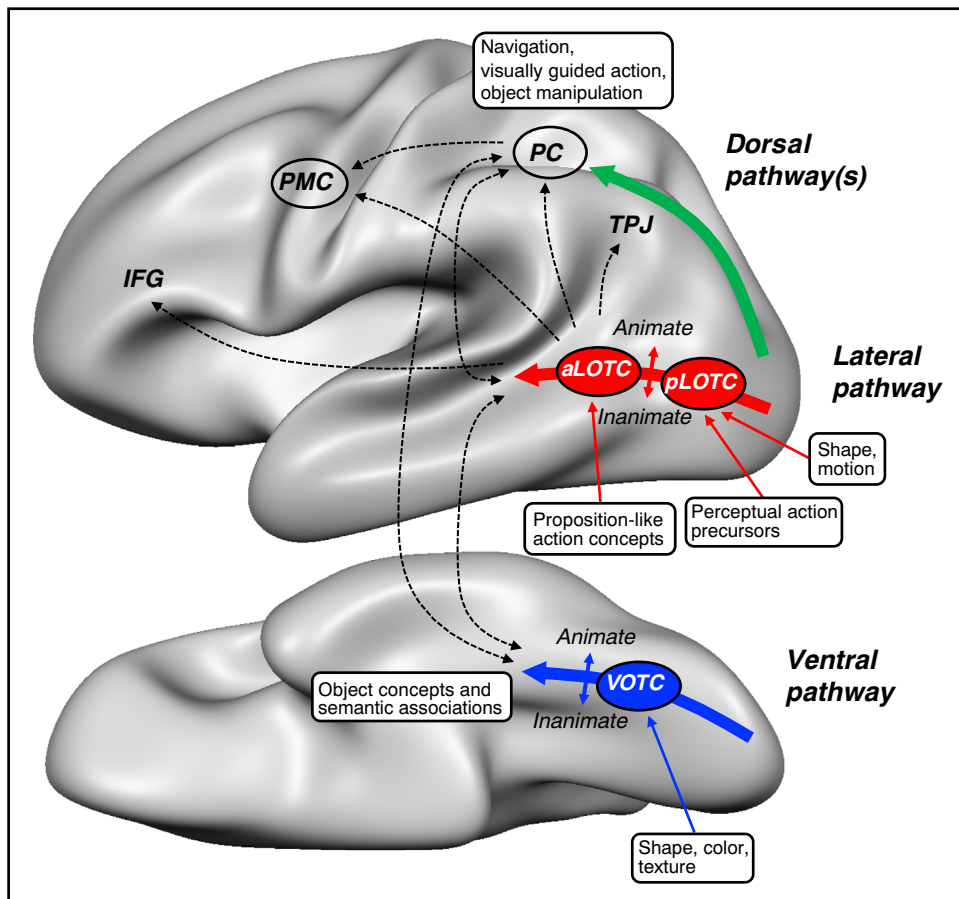
Figure 1. Overview of effects discussed. Numbers correspond to the peak coordinates of references listed in this article; letters refer to separate analyses/experiments within the same study. Labels of classical functionally defined areas for body parts, faces, and places are included for orientation. These labels are meant to serve as coarse landmarks rather than precise localizations of the functional areas. Reddish number/text colors indicate person/animate-related representations, blueish colors indicate inanimate-related representations. Colored areas indicate coarse outlines of specialized areas: person/animate-related (red), inanimate-related (blue), verbs (magenta), shape/form (yellow), conceptual/semantic action and object representations (white). Abbreviations: CoS, collateral sulcus; EBA, extrastriate body area; ErC, entorhinal cortex; FBA, fusiform body area; FFA, fusiform face area; FS, fusiform sulcus; hMT, human (middle temporal) motion area; IFG, lateral fusiform gyrus; LG, lingual gyrus; mFG, medial fusiform gyrus; OFA, occipital face area; OPA, occipital place area; OTS, occipitotemporal sulcus; PhC, parahippocampal cortex; PPA, parahippocampal place area; PrC, perirhinal cortex; RS, rhinal sulcus (see Figure 1 in Box 1 for lateral brain regions).

Distinct pathways for action and object recognition in the OTC

The representational organization in the LOTC can be described along two major dimensions: a posterior–anterior gradient from entities and perceptual action precursors to conceptual action representations and a dorsal–ventral distinction that segregates animate and inanimate action-relevant aspects. This organization suggests that the LOTC contains a hierarchically organized pathway for action recognition that integrates basic components of actions into conceptual, perhaps proposition-like, representations that capture the ‘what’ meaning of actions. This idea of an action pathway in the LOTC aligns well with previous, anatomically and functionally motivated proposals of a pathway in the LOTC [2, 10, 12–15]. However, there are some important extensions. Contrary to some previous proposals [13, 14], we propose that the lateral pathway processes not only motion but also other information that is critical for action recognition such as object shape and interobject relations. Moreover, we propose that the action pathway is organized along a major segregation – animate versus inanimate action aspects – that mirrors the organization of the VOTC (Figure 2, Key figure) (see Box 3 for a more detailed discussion of other proposals and their relations to the proposal made here).

Key figure

Schematic overview of the proposed organization of lateral, ventral, and dorsal pathways



Trends In Cognitive Sciences

Figure 2. Black broken arrows indicate functional connections between pathways and functional connections from pathways to other brain regions discussed in this article. Connections between pathways are displayed as double-sided arrows to indicate that flow of information is expected to be bidirectional without following a strict hierarchical order (lateral connections). Single-sided arrows indicate projections from pathways to downstream targets. Neither type of arrows implies directionality (all connections are expected to be reciprocal). The displayed connections do not constitute an exhaustive summary of all connections; connections beyond the scope of this article are not included. PMc and PC are highlighted to indicate their preeminent connectedness with LOTC. Abbreviations: a/pLOTC, anterior/posterior LOTC; IFG, inferior frontal gyrus; PC, parietal cortex; PMc, premotor cortex; TPJ, temporoparietal junction.

The proposal of a lateral action recognition pathway in the LOTC also has important implications for the interpretation of dorsal and ventral pathways. First, it suggests that action recognition – from perceptual to conceptual action representation – primarily draws on the LOTC rather than dorsal pathway regions in the frontoparietal cortex. This does not rule out that frontoparietal regions are involved in action recognition. Coactivation of the LOTC and frontoparietal regions during action recognition and the tight structural interconnections between these areas [52] point toward a strong interplay between these regions during action recognition. However, the

Box 3. An updated model of pathways in high-level visual cortex

The action recognition hierarchy in the LOTC coincides with a lateral structurally defined pathway that is distinct from dorsal and ventral pathways [2,13–15]. What are the similarities and differences with other proposals about the function of this pathway [10,12–14]?

Early proposals emphasized visual motion as the dominant feature processed in the lateral pathway [14]. Relatedly, Pitcher and Ungerleider highlight the sensitivity to moving faces and bodies and suggest that this pathway processes dynamic aspects of social perception [13] that do not match the ‘what’ and ‘where/how’ profiles of the ventral and dorsal pathways, respectively. Lingnau and Downing [10] highlight the diverse nature of action-related representations in the LOTC, including the representation of tools, verbs, and action concepts. They characterize the LOTC as a mosaic of several overlapping types of representations, some of which are organized along gradients (e.g., concrete–abstract, visual–multimodal, perceptual–intentional, and body–tool motion). Weiner and Grill-Spector [12] propose a lateral pathway from visual (e.g., form and motion) to multimodal processing (e.g., visual, tactile, and verbal). While not explicitly framed as a pathway for action recognition, they highlight the role of anterior LOTC as a convergence zone for action representation, similar to Lingnau and Downing’s proposal.

Our model (see [Figure 2](#) in main text) extends motion-centric models by considering also nonmotion aspects such as object shape and interobject relations. The models proposed by Weiner and Grill-Spector and Lingnau and Downing, and our model all ascribe action representation as a central function of the LOTC. While Lingnau and Downing offer a number of organizational axes, Weiner and Grill-Spector more explicitly propose a distinction between the posterior and anterior LOTC for perceptual versus modality-general representation that is similar to the posterior–anterior axis of organization in our model. A key difference with other lateral pathway models is that ours introduces a second organizational axis for animate and inanimate object and action aspects. A second key difference concerns the roles of, and the relation between, the lateral and ventral pathways. Previous accounts conceive the lateral pathway as being functionally distinct from both the ventral and dorsal pathways. However, there are significant parallels between the lateral and ventral pathways. Both pathways are similarly organized, with a major segregation between animate and inanimate action/object information. This organizational similarity might reflect similar, complementary functions in recognizing (and knowing about) actions and objects, respectively. In our model, lateral and ventral pathways subserve common superordinate functions in recognition and memory as opposed to the functionally more distinct dorsal ‘where/how’ pathway.

precise role of frontoparietal regions in this network remains unclear, and the evidence for different network models of action recognition is mixed [53]. Recent studies showed that occipitotemporal and frontoparietal regions form separate networks based on the distinct representational similarities of observed actions [54] and objects [55] within those regions. This seems to argue against a strict processing hierarchy (e.g., from the LOTC over parietal to frontal cortex) and rather suggests that occipitotemporal and frontoparietal regions provide different functions rather than a common function. Potential roles of frontoparietal regions during action recognition are discussed in [Box 2](#).

Second, our proposal sheds new light on the role of object representation in the LOTC and VOTC. Regions that preferentially respond to object categories (like body parts, faces, and tools) and a large-scale distinction into animate and inanimate entities are found in both the LOTC and VOTC [4,5] ([Figure 1](#)). These regions are typically discussed with a focus on object recognition and memory, and often they are considered to belong to the ventral visual system (as opposed to the dorsal pathway) [4,56]. A dominant proposal for the duplication of domain-specific response preferences is that the LOTC and VOTC process different aspects of objects related to motion and form, respectively, for example, how tools or animals typically move versus their shapes [41,42,48,57,58]. However, the evidence presented here suggests that the LOTC represents information beyond motion: LOTC action representations generalize across different movements, and the LOTC discriminates between two identical movements that are directed at different targets, such as objects and persons. In line with these findings, the LOTC also shows sensitivity to the form of objects [55,59,60], even for meaningless shapes that do not imply motion [61–64]. Finally, the anterior LOTC shows the same level of activation for verbs describing nonmotor, mental actions as for physical actions [36]. Our proposal offers a simple alternative to the form versus motion interpretation, namely that LOTC object representations

do not subserve object (feature) recognition as such but rather function as precursors for action recognition – as agents, instruments, and recipients of actions. By contrast, object recognition predominantly depends on the VOTC in the ventral pathway. This interpretation plausibly explains the duplication of object responses in the VOTC and LOTC. The VOTC is tuned to certain object domains (or features that correlate with a domain) in the service to recognize what object it is, whereas the LOTC is tuned to certain object domains/features in the service to recognize what the object is doing. Notably, some object features might be important for both action and object recognition (e.g., shape), whereas other features may be more involved with one or the other function (e.g., motion, color, and texture). Likewise, some object categories or object parts might be disproportionately relevant for either action or object recognition (e.g., tools, scenes, or certain body effectors, face parts, etc.). Predictions derived from these considerations are discussed in the following sections.

Dissociations of object aspects in the LOTC and VOTC

Motion, which is disproportionately important for action as compared to object recognition, activates the LOTC to a stronger degree than the VOTC [41,42]. By contrast, the reverse pattern is observed for color [65,66] and texture/material [61,63,64] – features that are important for object recognition (as well for motorically interacting with objects) but virtually irrelevant for action recognition. These distinctions may not be clear cut; for example, subtle color responses have been found in an LOTC region that seems consistent with the homologue macaque area V4d [67], which also responds to color [68]. Both the LOTC and VOTC are sensitive to object shape [61,62], which is important for both action and object recognition.

The LOTC responds more strongly to predominantly action-involved body parts (hands and feet) relative to noneffector body parts like waist and chest, whereas the reverse pattern is found in the VOTC [69]. Moreover, posterior superior temporal sulcus (pSTS) appears more sensitive to facial actions and expressions [70,71], whereas the fusiform face area (FFA) appears more sensitive to face identity [72–74], although this dissociation seems not absolute [75,76]. Likewise, the VOTC is more sensitive than the LOTC in discriminating animal types (e.g., *primates*, *reptiles*, and *birds*), whereas the LOTC is more sensitive than the VOTC in discriminating actions of animals (e.g., *fighting*, *eating*, and *running*) [77].

Representations in lateral areas for faces, human bodies, and inanimate objects are more part based and orientation dependent, whereas representations in their ventral counterparts are more holistic and dependent on correct configurations of their parts [78]. This fits the proposal made here since recognizing actions relies more on specific object parts (e.g., *hands*, *feet*, *handles*, and *knobs*), whereas object recognition requires capturing objects as wholes. Viewpoint-invariant representations of body part postures, which are indicative of actions like grasping or communicative gestures, have been found in the human LOTC, but not VOTC [79].

Tools are exclusively made to support actions. Hence, recognizing a tool usually serves to recognize its function, and to act with it. According to our proposal, the LOTC encodes the action associated with a tool, that is, its function, whereas the VOTC encodes features important for using the tool, for example, size and weight (e.g., *scissors* versus *shears*) and texture and material (*sharp* and *metal*). Naming the functions of tools activates the LOTC [80], and the LOTC responds to tool words also in blind individuals, suggesting that tool representations in the LOTC are accessible independently of visual shape or motion [81]. Studies specifically investigating surface properties of tools are sparse. Surface properties of tools are not only relevant for manipulating a tool but also for its function. Thus, tool function often covaries with certain properties (e.g., that both knives and scissors have a sharp metal part), which might explain why surface-property-sensitive regions in the VOTC group tools based on their functions [82,83].

Finally, dissociations are also observed for scenes in line with our proposal. The parahippocampal place area (PPA) in the VOTC seems most sensitive to scene layout and landmarks [7,84], whereas in the LOTC, scenes are organized along actions that usually take place in them [85]. This scene effect in the LOTC does not coincide with the occipital place area (OPA), which is located more posteriorly [86] and might be better ascribed to the dorsal pathway, although anterior parts of the OPA might overlap with the LOTC as defined here. Although the OPA is more strongly associated with the detection of scene boundaries and navigational affordances [7], TMS to the (right) OPA impairs scene discrimination [87,88]; a function that is usually not associated with the how/where profile of the dorsal pathway [89].

Relationship between functional topography and retinotopy

Many of the functional responses described in the previous section overlap with retinotopic maps in the lateral and ventral pathways. How is action- and object-related information topographically related to retinotopic organization? In the LOTC, retinotopic maps appear to overlap the posterior LOTC, but not anterior LOTC [86,90]. By contrast, in the VOTC, some anterior regions, for example, the PPA, show retinotopic organization, but retinotopic maps seem restricted to regions medial to the fusiform sulcus [91,92], whereas lateral areas like the FFA do not overlap known retinotopic maps and are foveally biased [93]. A reason for this organization could be that medial VOTC regions have preferences for position-specific object information important for navigation and object manipulation, whereas the lateral VOTC and anterior LOTC preferentially process less position-relevant information like animate entities and their actions, respectively.

Another noteworthy distinction, revealed by human fMRI [94] and macaque tracer studies [95], is that the anterior VOTC shows a preference for the upper visual field, whereas a reverse preference for the lower visual field is found in the LOTC [56]. Motion-responsive regions, like the MT, MST, and pSTS, do not show this visual field bias [96]. This seems to imply that object recognition also relies on input from lateral visual cortex. Lateral and ventral (and likely also dorsal) pathways thus might be less segregated at their origins and become increasingly specialized in anterior regions. In line with this view, TMS-induced disruption to lateral occipital regions showing preferences for faces (OFA), body parts (EBA), scenes (OPA), and inanimate objects (LO) result in category-selective recognition interference [88,97], suggesting that object recognition depends on posterior LOTC, at least partially.

Dissociation of higher-level action and object representations in the LOTC and VOTC

In the previous section, we focused on relatively basic, perceptual representations in the LOTC and VOTC. The organization of more abstract, semantic knowledge is less well understood. However, there is evidence that also more anterior temporal regions show preferences for either object- or action-related information. For example, object representations that generalize across stimulus modality are found in the anterior VOTC [98,99], roughly parallel to modality-general action representations in the anterior LOTC. More anteriorly in the ventral temporal cortex, the perirhinal cortex is sensitive to semantic features that are relevant for basic level object discrimination (e.g., zebras but not horses have stripes) [100–102]. By contrast, the left lateral mid-temporal cortex differentiates the roles that objects play as either agents or recipients in action descriptions (e.g., *man chases dog* vs *dog chases man*) [103], which might point toward a more integrated representation of specific actions/events (*who did what to whom*). Thus, taken together, ventral and lateral pathways reveal roughly comparable gradients of representation: from perceptual features (duplicate if relevant for both action and object recognition) to superordinate categorical distinctions (e.g., person/object-directedness, animacy) to more specific semantic representations of actions and objects.

Connectivity between pathways and lateralization

In the previous sections, we proposed that lateral and ventral pathways subserve similar, but complementary functions in action and object recognition, respectively. How do these pathways interact with each other and with the dorsal pathway as well as other brain regions? There is evidence for different networks that connect subregions in lateral, ventral, and/or dorsal pathways to subserve common functions. A prominent example is object manipulation and tool use, which draws on a network that connects VOTC regions representing object surface properties (e.g., *graphite and rubber endings of a pencil*), LOTC regions representing tool function (*writing, erasing*), and superior and inferior parietal regions for correctly grasping and manipulating an object, respectively [82,104–106]. As another example, scene-selective regions in the VOTC form a more medial network with superior parietal regions (and precuneus) to subserve navigation [107]. These interpathway connections (among others) argue against a strict segregation of pathways. In addition, the different pathways may form the basis for various downstream functions. For example, proposition-like action representations in the anterior LOTC might form the basis for language-related functions in superior temporal and inferior frontal cortex [108]. Body- and face-related representations in posterior LOTC may form the basis for representing others' mental states like feelings, desires, and intentions in regions of the mentalizing network such as the temporoparietal junction (TPJ) [109].

Several aspects about action representation are left-lateralized, such as tools and hands [5,6,40,47] as well as modality-general action representations [27]. This lateralization likely reflects associations to left-lateralized networks, for example, for tool use [82,104–106] and language [108]. By contrast, areas sensitive to body- and face-related cues such as the FFA, EBA, and pSTS [13,43], but also mentalizing-related brain regions such as TPJ [109], are more pronounced in the right hemisphere. A plausible interpretation of these hemispheric dissociations is that action recognition and representation (what someone is doing, what can be done with an object, etc.) primarily draws on the left LOTC, whereas the processing of body-related cues that are important for other aspects of social perception – such as the recognition of emotional expressions, eye gaze, etc. – which is critical for representing others' mental states, primarily draws on the right LOTC. However, observed actions can be decoded equally well in left and right hemispheres, suggesting that the right LOTC is also important for action recognition [24,25]. How the left and right LOTC exchange information during action observation and their exact roles in action-related representation remain a challenge for future research [110].

Concluding remarks

The evidence that the LOTC plays a central role in action recognition invites rethinking the general functions of pathways in posterior cortex. Here, we build on the original conception of the ventral stream as a 'what' pathway and propose that ventral and lateral OTC are specialized for distinct aspects of recognition: to recognize objects and what these objects do, respectively. This model is certainly an idealization and meant to emphasize major organizational principles – it remains to be determined how strongly the functions of lateral and ventral pathways dissociate and where and how they interact. Another pressing question concerns the interplay between LOTC and frontoparietal regions during action (and object) recognition. Addressing these and other questions that follow from our framework (see [Outstanding questions](#)) may help advance our understanding of knowledge organization in the brain.

Acknowledgments

We thank Yanchao Bi, Angelika Lingnau, Marius Peelen, and Gilles Vannuscorps for their critical comments on an earlier version of this manuscript. The publication of this work was supported by the Ministero dell'Università e della Ricerca project 'Dipartimenti di Eccellenza'.

Outstanding questions

What is the role of frontoparietal cortex in action recognition and semantic processing? To what extent and how are frontoparietal areas also involved in object recognition?

How strictly is action recognition left-lateralized and how do the left and right LOTC interact?

Which parts of the brain subserve the integration of action concepts (e.g., *opening* independently of who is opening and what is opened) with object knowledge (e.g., to understand that *Mary opens a present*)?

Are physical actions that (directly) induce a change of states in the world represented in the same system as nonphysical actions (e.g., mental actions and perception actions)? How are animal actions (e.g., the biting of a mosquito) and inanimate events (e.g., a rock rolling down a hill) represented in the brain?

Where is conceptual knowledge about actions' underlying motives and intentions that are not directly perceptible (e.g., drinking to *quench thirst*) represented in the brain and how does it relate to the representation of actions?

Does the auditory ventral stream also subdivide into substreams for action and object recognition? How do the organizations of auditory and visual ventral streams relate to each other?

Declaration of interests

The authors declare no conflicts of interest.

References

1. Ungerleider, L.G. and Mishkin, M. (1982) Two cortical visual systems. In *Analysis of Visual Behavior* (Ingle, D.J. et al., eds), MIT Press
2. Goodale, M.A. and Milner, A.D. (1992) Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25
3. Weiner, K.S. et al. (2017) The cytoarchitecture of domain-specific regions in human high-level visual cortex. *Cereb. Cortex* 27, 146–161
4. Konkle, T. and Caramazza, A. (2013) Tripartite organization of the ventral stream by animacy and object size. *J. Neurosci.* 33, 10235–10242
5. Chao, L.L. et al. (1999) Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat. Neurosci.* 2, 913–919
6. Bracci, S. and Peelen, M.V. (2013) Body and object effectors: the organization of object representations in high-level visual cortex reflects body-object interactions. *J. Neurosci.* 33, 18247–18258
7. Epstein, R.A. and Baker, C.I. (2019) Scene perception in the human brain. *Annu. Rev. Vis. Sci.* 5, 373–397
8. Mahon, B.Z. and Caramazza, A. (2011) What drives the organization of object knowledge in the brain? *Trends Cogn. Sci.* 15, 97–103
9. Peelen, M.V. and Downing, P.E. (2017) Category selectivity in human visual cortex: beyond visual object recognition. *Neuropsychologia* 105, 177–183
10. Lingnau, A. and Downing, P.E. (2015) The lateral occipitotemporal cortex in action. *Trends Cogn. Sci.* 19, 268–277
11. Leshinskaya, A. et al. (2020) Concepts of actions and their objects. In *The Cognitive Neurosciences* (Gazzaniga, M. et al., eds), pp. 757–765, MIT Press
12. Weiner, K.S. and Grill-Spector, K. (2013) Neural representations of faces and limbs neighbor in human high-level visual cortex: evidence for a new organization principle. *Psychol. Res.* 77, 74–97
13. Pitcher, D. and Ungerleider, L.G. (2021) Evidence for a third visual pathway specialized for social perception. *Trends Cogn. Sci.* 25, 100–110
14. Boussaoud, D. et al. (1990) Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *J. Comp. Neurol.* 296, 462–495
15. Haak, K.V. and Beckmann, C.F. (2018) Objective analysis of the topological organization of the human cortical visual connectome suggests three visual pathways. *Cortex* 98, 73–83
16. Oosterhof, N.N. et al. (2013) Crossmodal and action-specific: neuroimaging the human mirror neuron system. *Trends Cogn. Sci.* 17, 311–318
17. Orban, G.A. et al. (2021) From Observed Action Identity to Social Affordances. *Trends Cogn. Sci.* 25, 493–505
18. Caramazza, A. et al. (2014) Embodied cognition and mirror neurons: a critical assessment. *Annu. Rev. Neurosci.* 37, 1–15
19. Tarhan, L.Y. et al. (2015) Shared and distinct neuroanatomic regions critical for tool-related action production and recognition: evidence from 131 left-hemisphere stroke patients. *J. Cogn. Neurosci.* 27, 2491–2511
20. Vannuscorps, G. et al. (2016) Persistent sparing of action conceptual processing in spite of increasing disorders of action production: a case against motor embodiment of action concepts. *Cogn. Neuropsychol.* 33, 191–219
21. Urgesi, C. et al. (2014) Neuroanatomical substrates of action perception and understanding: an anatomic likelihood estimation meta-analysis of lesion-symptom mapping studies in brain injured patients. *Front. Hum. Neurosci.* 8, 344
22. Vannuscorps, G. et al. (2019) Large-scale organization of the hand action observation network in individuals born without hands. *Cereb. Cortex* 29, 3434–3444
23. Wurm, M.F. et al. (2016) Decoding concrete and abstract action representations during explicit and implicit conceptual processing. *Cereb. Cortex* 26, 3390–3401
24. Wurm, M.F. et al. (2017) Action categories in lateral occipitotemporal cortex are organized along sociality and transitivity. *J. Neurosci.* 37, 562–575
25. Wurm, M.F. and Lingnau, A. (2015) Decoding actions at different levels of abstraction. *J. Neurosci.* 35, 7727–7735
26. Hafri, A. et al. (2017) Neural representations of observed actions generalize across static and dynamic visual input. *J. Neurosci.* 37, 3056–3071
27. Wurm, M.F. and Caramazza, A. (2019) Distinct roles of temporal and frontoparietal cortex in representing actions across vision and language. *Nat. Commun.* 10, 289
28. Tarhan, L. and Konkle, T. (2019) Sociality and interaction envelope organize visual action representations. *Nat. Commun.* 11, 3002
29. Wurm, M.F. and Caramazza, A. (2019) Lateral occipitotemporal cortex encodes perceptual components of social actions rather than abstract representations of sociality. *Neuroimage* 202, 116153
30. Tucciarelli, R. et al. (2019) The representational space of observed actions. *eLife* 8, e47686
31. Afkalo, T. et al. (2020) A shared neural substrate for action verbs and observed actions in human posterior parietal cortex. *Sci. Adv.* 6, eabb3984
32. Lanzilotto, M. et al. (2020) Stable readout of observed actions from format-dependent activity of monkey's anterior intraparietal neurons. *Proc. Natl. Acad. Sci. U. S. A.* 117, 16596–16605
33. Papeo, L. et al. (2015) The origin of word-related motor activity. *Cereb. Cortex* 25, 1668–1675
34. Perrett, D.I. et al. (1989) Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.* 146, 87–113
35. Watson, C.E. et al. (2013) Action concepts in the brain: an activation likelihood estimation meta-analysis. *J. Cogn. Neurosci.* 25, 1191–1205
36. Bedny, M. et al. (2008) Concepts are more than percepts: the case of action verbs. *J. Neurosci.* 28, 11347–11353
37. Peelen, M.V. et al. (2012) Independent representations of verbs and actions in left lateral temporal cortex. *J. Cogn. Neurosci.* 24, 2096–2107
38. Hernandez, M. et al. (2014) Predication drives verb cortical signatures. *J. Cogn. Neurosci.* 26, 1829–1839
39. Downing, P.E. et al. (2001) A cortical area selective for visual processing of the human body. *Science* 293, 2470–2473
40. Striem-Amit, E. et al. (2017) Sensorimotor-independent development of hands and tools selectivity in the visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 114, 4787–4792
41. Beauchamp, M.S. et al. (2002) Parallel visual motion processing streams for manipulable objects and human movements. *Neuron* 34, 149–159
42. Beauchamp, M.S. et al. (2003) fMRI responses to video and point-light displays of moving humans and manipulable objects. *J. Cogn. Neurosci.* 15, 991–1001
43. Han, Z. et al. (2013) Distinct regions of right temporal cortex are associated with biological and human-agent motion: functional magnetic resonance imaging and neuropsychological evidence. *J. Neurosci.* 33, 15442–15453
44. Abassi, E. and Papeo, L. (2020) The representation of two-body shapes in the human visual cortex. *J. Neurosci.* 40, 852–863
45. Quadflieg, S. et al. (2015) The neural basis of perceiving person interactions. *Cortex* 70, 5–20
46. Baldassano, C. et al. (2017) Human-object interactions are more than the sum of their parts. *Cereb. Cortex* 27, 2276–2288

47. Bracci, S. *et al.* (2012) Closely overlapping responses to tools and hands in left lateral occipitotemporal cortex. *J. Neurophysiol.* 107, 1443–1456
48. Martin, A. and Weisberg, J. (2003) Neural foundations for understanding social and mechanical concepts. *Cogn. Neuropsychol.* 20, 575–587
49. He, C. *et al.* (2020) Roles of category, shape, and spatial frequency in shaping animal- and tool-selectivity in the occipitotemporal cortex. *J. Neurosci.* 40, 5644–5657
50. Isik, L. *et al.* (2017) Perceiving social interactions in the posterior superior temporal sulcus. *Proc. Natl. Acad. Sci. U. S. A.* 114, E9145–E9152
51. Papeo, L. *et al.* (2019) The large-scale organization of gestures and words in the middle temporal gyrus. *J. Neurosci.* 39, 5966–5974
52. Nelissen, K. *et al.* (2011) Action observation circuits in the macaque monkey cortex. *J. Neurosci.* 31, 3743–3756
53. Kemmerer, D. (2021) What modulates the mirror neuron system during action observation?: Multiple factors involving the action, the actor, the observer, the relationship between actor and observer, and the context. *Progr. Neurobiol.* 205, 102128
54. Yargholi, E. *et al.* (2021) Two distinct networks containing position-tolerant representations of actions in the human brain. *bioRxiv* Published online August 4, 2021. <https://doi.org/10.1101/2021.06.17.448825>
55. Bracci, S. and Op de Beeck, H. (2016) Dissociations and associations between shape and category representations in the two visual pathways. *J. Neurosci.* 36, 432–444
56. Kravitz, D.J. *et al.* (2013) The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* 17, 26–49
57. Martin, A. (2016) GRAPES-Grounding representations in action, perception, and emotion systems: How object properties and categories are represented in the human brain. *Psychon. Bull. Rev.* 23, 979–990
58. Chatterjee, A. (2010) Disembodying cognition. *Lang. Cogn.* 2, 79–116
59. Proklova, D. *et al.* (2016) Disentangling representations of object shape and object category in human visual cortex: the animate-inanimate distinction. *J. Cogn. Neurosci.* 28, 680–692
60. Wang, X. *et al.* (2018) Disentangling representations of shape and action components in the tool network. *Neuropsychologia* 117, 199–210
61. Cavina-Pratesi, C. *et al.* (2010) Separate channels for processing form, texture, and color: evidence from fMRI adaptation and visual object agnosia. *Cereb. Cortex* 20, 2319–2332
62. Peuskens, H. *et al.* (2004) Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. *J. Cogn. Neurosci.* 16, 665–682
63. Cant, J.S. *et al.* (2009) fMRI-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. *Exp. Brain Res.* 192, 391–405
64. Cant, J.S. and Goodale, M.A. (2007) Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb. Cortex* 17, 713–731
65. Fan, S. *et al.* (2020) Visual featural topography in the human ventral visual pathway. *J. Vis.* 20, 1029
66. Bartels, A. and Zeki, S. (2000) The architecture of the colour centre in the human visual brain: new results and a review. *Eur. J. Neurosci.* 12, 172–193
67. Wade, A. *et al.* (2008) fMRI measurements of color in macaque and human. *J. Vis.* 8, 6 1–19
68. Conway, B.R. *et al.* (2007) Specialized color modules in macaque extrastriate cortex. *Neuron* 56, 560–573
69. Bracci, S. *et al.* (2015) Representational similarity of body parts in human occipitotemporal cortex. *J. Neurosci.* 35, 12977–12985
70. Deen, B. and Saxe, R. (2019) Parts-based representations of perceived face movements in the superior temporal sulcus. *Hum. Brain Mapp.* 40, 2499–2510
71. Deen, B. *et al.* (2020) Processing communicative facial and vocal cues in the superior temporal sulcus. *Neuroimage* 221, 117191
72. Zhang, H. *et al.* (2016) Face-selective regions differ in their ability to classify facial expressions. *Neuroimage* 130, 77–90
73. Calder, A.J. and Young, A.W. (2005) Understanding the recognition of facial identity and facial expression. *Nat. Rev. Neurosci.* 6, 641–651
74. Tsantani, M. *et al.* (2021) FFA and OFA encode distinct types of face identity information. *J. Neurosci.* 41, 1952–1969
75. Anzellotti, S. and Caramazza, A. (2017) Multimodal representations of person identity individuated with fMRI. *Cortex* 89, 85–97
76. Bernstein, M. and Yovel, G. (2015) Two neural pathways of face processing: a critical evaluation of current models. *Neurosci. Biobehav. Rev.* 55, 536–546
77. Nastase, S.A. *et al.* (2017) Attention selectively reshapes the geometry of distributed semantic representation. *Cereb. Cortex* 27, 4277–4291
78. Taylor, J.C. and Downing, P.E. (2011) Division of labor between lateral and ventral extrastriate representations of faces, bodies, and objects. *J. Cogn. Neurosci.* 23, 4122–4137
79. Bracci, S. *et al.* (2018) View-invariant representation of hand postures in the human lateral occipitotemporal cortex. *Neuroimage* 181, 446–452
80. Martin, A. *et al.* (1995) Discrete cortical regions associated with knowledge of color and knowledge of action. *Science* 270, 102–105
81. Peelen, M.V. *et al.* (2013) Tool selectivity in left occipitotemporal cortex develops without vision. *J. Cogn. Neurosci.* 25, 1225–1234
82. Garcea, F.E. *et al.* (2019) Domain-specific diaschisis: lesions to parietal action areas modulate neural responses to tools in the ventral stream. *Cereb. Cortex* 29, 3168–3181
83. Chen, Q. *et al.* (2018) Abstract representations of object-directed action in the left inferior parietal lobule. *Cereb. Cortex* 28, 2162–2174
84. Malcolm, G.L. *et al.* (2016) Making sense of real-world scenes. *Trends Cogn. Sci.* 20, 843–856
85. Groen, I.I. *et al.* (2018) Distinct contributions of functional and deep neural network features to representational similarity of scenes in human brain and behavior. *eLife* 7, e32962
86. Silson, E.H. *et al.* (2015) A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *J. Neurosci.* 35, 11921–11935
87. Dilks, D.D. *et al.* (2013) The occipital place area is causally and selectively involved in scene perception. *J. Neurosci.* 33, 1331–1336
88. Wischniewski, M. and Peelen, M.V. (2021) Causal evidence for a double dissociation between object- and scene-selective regions of visual cortex: a preregistered TMS replication study. *J. Neurosci.* 41, 751–756
89. Freud, E. *et al.* (2016) 'What' is happening in the dorsal visual pathway. *Trends Cogn. Sci.* 20, 773–784
90. Larsson, J. and Heeger, D.J. (2006) Two retinotopic visual areas in human lateral occipital cortex. *J. Neurosci.* 26, 13128–13142
91. Arcaro, M.J. *et al.* (2009) Retinotopic organization of human ventral visual cortex. *J. Neurosci.* 29, 10638–10652
92. Arcaro, M.J. and Livingstone, M.S. (2017) Retinotopic organization of scene areas in macaque inferior temporal cortex. *J. Neurosci.* 37, 7373–7389
93. Grill-Spector, K. and Weiner, K.S. (2014) The functional architecture of the ventral temporal cortex and its role in categorization. *Nat. Rev. Neurosci.* 15, 536–548
94. Kravitz, D.J. *et al.* (2010) High-level visual object representations are constrained by position. *Cereb. Cortex* 20, 2916–2925
95. Ungerleider, L.G. *et al.* (2008) Cortical connections of area V4 in the macaque. *Cereb. Cortex* 18, 477–499
96. Huk, A.C. *et al.* (2002) Retinotopy and functional subdivision of human areas MT and MST. *J. Neurosci.* 22, 7195–7205
97. Pitcher, D. *et al.* (2009) Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Curr. Biol.* 19, 319–324
98. Fairhall, S.L. and Caramazza, A. (2013) Brain regions that represent amodal conceptual knowledge. *J. Neurosci.* 33, 10552–10558
99. Simanova, I. *et al.* (2014) Modality-independent decoding of semantic information from the human brain. *Cereb. Cortex* 24, 426–434

100. Clarke, A. and Tyler, L.K. (2014) Object-specific semantic coding in human perirhinal cortex. *J. Neurosci.* 34, 4766–4775
101. Bruffaerts, R. et al. (2013) Similarity of fMRI activity patterns in left perirhinal cortex reflects semantic similarity between words. *J. Neurosci.* 33, 18597–18607
102. Martin, C.B. et al. (2018) Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. *eLife* 7, e31873
103. Frankland, S.M. and Greene, J.D. (2015) An architecture for encoding sentence meaning in left mid-superior temporal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 112, 11732–11737
104. Goldenberg, G. and Spatt, J. (2009) The neural basis of tool use. *Brain* 132, 1645–1655
105. Binkofski, F. and Buxbaum, L.J. (2013) Two action systems in the human brain. *Brain Lang.* 127, 222–229
106. Rizzolatti, G. and Matelli, M. (2003) Two different streams form the dorsal visual system: anatomy and functions. *Exp. Brain Res.* 153, 146–157
107. Kravitz, D.J. et al. (2011) A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.* 12, 217–230
108. Hickok, G. and Poeppel, D. (2007) The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393
109. Carter, R.M. and Huettel, S.A. (2013) A nexus model of the temporal-parietal junction. *Trends Cogn. Sci.* 17, 328–336
110. Akbiyik, S. et al. (2021) Shared and distinct neural representations of human-agent actions and inanimate events. *J. Vis.* 21, 2490
111. Thompson, E.L. et al. (2019) Conceptualizing and testing action understanding. *Neurosci. Biobehav. Rev.* 105, 106–114
112. Vannuscorps, G. and Caramazza, A. (2016) Typical action perception and interpretation without motor simulation. *Proc. Natl. Acad. Sci. U. S. A.* 113, 86–91
113. Hickok, G. (2014) *The Myth of Mirror Neurons: the Real Neuroscience of Communication and Cognition*, W.W. Norton
114. Heyes, C. and Catmur, C. (2021) What happened to mirror neurons? *Perspect. Psychol. Sci.* Published online July 9, 2021. <https://doi.org/10.1177/1745691621990638>
115. Buxbaum, L.J. and Kalenine, S. (2010) Action knowledge, visuomotor activation, and embodiment in the two action systems. *Ann. N. Y. Acad. Sci.* 1191, 201–218
116. Leshinskaya, A. and Caramazza, A. (2014) Nonmotor aspects of action concepts. *J. Cogn. Neurosci.* 26, 2863–2879
117. Spunt, R.P. and Lieberman, M.D. (2012) Dissociating modality-specific and supramodal neural systems for action understanding. *J. Neurosci.* 32, 3575–3583
118. Thornton, M.A. and Mitchell, J.P. (2018) Theories of person perception predict patterns of neural activity during mentalizing. *Cereb. Cortex* 28, 3505–3520
119. Vannuscorps, G. and Caramazza, A. (2016) Impaired short-term memory for hand postures in individuals born without hands. *Cortex* 83, 136–138
120. Wurm, M.F. et al. (2014) Predicting goals in action episodes attenuates BOLD response in inferior frontal and occipitotemporal cortex. *Behav. Brain Res.* 274, 108–117