

Chapter 13

Domain-specific connectivity drives the organization of object knowledge in the brain

BRADFORD Z. MAHON*

Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, United States

Abstract

The goal of this chapter is to review neuropsychological and functional MRI findings that inform a theory of the causes of functional specialization for semantic categories within occipito-temporal cortex—the ventral visual processing pathway. The occipito-temporal pathway supports visual object processing and recognition. The theoretical framework that drives this review considers visual object recognition through the lens of how “downstream” systems interact with the outputs of visual recognition processes. Those downstream processes include conceptual interpretation, grasping and object use, navigating and orienting in an environment, physical reasoning about the world, and inferring future actions and the inner mental states of agents. The core argument of this chapter is that innately constrained connectivity between occipito-temporal areas and other regions of the brain is the basis for the emergence of neural specificity for a limited number of semantic domains in the brain.

INTRODUCTION

Defining the question

The goal of this chapter is to review neuropsychological and functional MRI findings that inform a theory of the causes of functional specialization for semantic categories within occipito-temporal areas that support visual object processing and recognition. The theoretical framework that drives this review considers visual object recognition through the lens of how “downstream” systems interact with the outputs of recognition processes. Those downstream processes include conceptual interpretation, grasping and object use, navigating and orienting in an environment, physical reasoning about the world, and inferring future actions and the inner mental states of agents (Johnson et al., 1991; Spelke and Kinzler, 2007). The core argument of this chapter is that innately constrained domain-specific connectivity between occipito-temporal areas and other regions of the brain is the basis for the emergence of neural specificity for

a limited number of semantic domains in occipito-temporal cortex (Mahon et al., 2007, 2009; Martin, 2007; Mahon and Caramazza, 2009, 2011).

The domain-specific hypothesis that motivates this review can be separated into two claims: (i) connectivity provides the initial scaffolding for the organization of occipito-temporal areas by semantic domain and (ii) that connectivity is largely hardwired (i.e., innately constrained). On this view, connectivity constrains which types of computations are represented in which areas of occipito-temporal cortex. Experience must, of course, provide the content that is represented.

As will be developed, by “domains” I do not mean exactly “semantic categories”—although some semantic categories may align with domains. Categories, on my deployment of the term, pick out groupings of things in the world. Domains are individuated by the different types of computations that are required to support (different) behavioral goals. Brain “regions” do not implement the types of computations that distinguish domains.

*Correspondence to: Bradford Z. Mahon, Ph.D., 327E Baker Hall, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, Pennsylvania 15213, United States. Tel: +1-412-268-7136, E-mail: bmahon@andrew.cmu.edu

Thus, brain regions are not domain-specific; a network of multiple regions is the smallest unit in the brain that can meaningfully be referred to as “domain-specific.” By hypothesis, category-preferring regions in occipito-temporal areas have those preferences because of how broader networks of regions are able to process their outputs in the service of computational goals.

Thus, domain-specific constraints are not “local” to the occipital-temporal cortex—although the occipital-temporal cortex is the primary place where such theories have been developed and tested. Domain-specific constraints are aligned with the computational goals required of successful processing of items from different domains (e.g., navigating vs inferring someone’s mental state). Alfonso Caramazza and I have referred to earlier iterations of this framework as the distributed domain-specific hypothesis (Mahon and Caramazza, 2011) to emphasize that a given domain-specific neural system is distributed over dissociable brain regions, where each region contributes to a larger computation that “defines” the domain.

Criteria for domain-specificity

Research on the functional organization of the temporal lobe can (roughly) be divided into “activation evidence” (e.g., fMRI, neurophysiology, and EEG) and “causal evidence” (e.g., human neuropsychological investigations, animal ablation studies, lesion-behavior correlation, electrical stimulation, and TMS). Studies using activation evidence to study the organization of occipito-temporal areas have traditionally emphasized a specific approach for testing the specificity of a brain region: systematically vary the types of stimuli presented, until one “type” of stimulus is found that maximally drives neural responses in the area in question. This can be done in a data-driven approach at first—show the subject a lot of different kinds of stimuli and see which ones drive the neuron or voxel. This approach may or may not use a threshold criterion to infer specificity—for instance, that the response to stimulus X is at least twice as large as the response to any other stimulus type, e.g., Downing et al. (2006).

Discoveries of regions that consistently exhibit category-preferences are naturally followed by studies that parametrically deconstruct the preferred stimulus type to understand what it is about that stimulus type that drives responses in the region. For instance, if a region responds to faces, then an approach would be to experimentally deconstruct faces to understand which aspects of a face are driving the region (is it the eyes, or the organization of the features of the face, and so on). Such studies have provided an incredibly rich basis for

understanding *how* occipito-temporal regions come to be active for a given category of stimuli. That broad paradigm has been reinforced by the practice of naming functionally defined subregions of the occipito-temporal cortex with the category-name of the stimulus type that elicits a maximal response: for instance, the “fusiform face area,” “parahippocampal place area,” “visual word form area,” “extra-striate body-part area,” and “the number form area.”

If a theory of what a region *does* amounts to a description of the types of situations in which it is maximally active, there is the risk that the data become the theory (Poeppel, 2012; Zhang et al., 2014). The “data” are that the fusiform face area is maximally active for faces compared to nonfaces; the theory is that the fusiform face is specialized for face processing. If another type of stimulus activates the fusiform face area, there are two broad choices for such a proposal. Either: Give up the idea that said the “face area” is face-specific; Or: Explain the new finding in terms of the similarity to faces of the nonface stimuli that drive the face area. For instance, greebles (nonface) stimuli were shown to also activate the face area. Why? One suggestion was that they look like faces (in such and such specific ways that may even be quantifiable). Such a response is sufficient as an account of *how* the FFA came to be active for a greeble; but it only distracts from the *why*. Characterizing the selectivity of a region, or lack thereof, does not resolve theoretical issues as to what the region does. Describing the selectivity of responses in a region is invaluable information—but it is just the starting point for constraining a theory of what the region does.

As another example: The construal of domain-specificity that argues that occipito-temporal regions are innately specialized for different categories has been challenged on the grounds that response profiles are not all or none when comparing the preferred to nonpreferred stimuli. If category selectivity of neural responses in a region is taken to be evidence *for* domain-specificity, then demonstrations of *nonselectivity* in the neural responses of that region should count as evidence against domain-specificity. For instance, using functional neuroimaging, a region might respond maximally to faces, and less so but still substantially to animals, and less so but still a bit to tools, and even less to places. The area is defined as the “face area”—and yet the response to nonfaces in the face area may be above a fixation baseline. Similarly, multivoxel pattern analyses indicate that substantial information about nonpreferred categories (e.g., tools and places) is present in regions that are (putatively) specialized for other categories (e.g., faces, animals; Haxby et al., 2001). Critics of domain-specificity are correct that demonstrations of “category-specific”

neural responses do not, ipso facto, constitute evidence for domain-specificity.

This chapter argues for an alternative view of domain-specificity. According to the distributed domain-specific view, domain-specificity reduces to the claim that there are innately constrained patterns of connectivity at the granularity of different computational problems that need to be solved. Domains are individuated by the divergent computational goals that must be solved, not the types of stimuli that drive responses in one or another brain region (Mahon and Caramazza, 2011; Conway, 2018). Such computational goals are supported by processing across multiple regions. Thus, “domain-specificity” describes the computational goals of a *network* of regions, rather than the stimulus preferences of any given region. It follows that the test of domain-specificity of a region does not reduce to a test of whether that region exhibits neural responses that are differential for one category of stimuli compared to others. The region may be category-specific in its response profile—but it may not be, while still being part of a domain-specific network. This view emphasizes that in order to understand the constraints that shape the organization of the temporal lobe, we need to look outside of the temporal lobe.

Scope of this review

Everyday recognition and use of manipulable objects involves the integration of visual object representations in the occipito-temporal cortex with object-directed action representations in parietal and frontal areas. Studying the neural organization of manipulable objects provides an instructive perspective on broader issues that are central to understanding the constraints that shape the organization of the occipito-temporal cortex. “Manipulable objects” refer to any object that can be grasped and manipulated (e.g., keys, fork, hat, smartphone). The empirical review of this chapter emphasizes the neural systems that support recognition, grasping, and use of manipulable objects (Fig. 13.1). The discussion will briefly sketch how the distributed domain-specific hypothesis applies to understanding functional specialization in occipital-temporal areas for faces, animals, written words, body parts, and places.

The scope of this empirical review is primarily limited to neuropsychological and functional MRI studies that inform theories of how visual categories are organized in the temporal lobe, and the constraints that lead to a *consistent* organization across individuals in occipito-temporal areas. The empirical review does not cover development (e.g., Johnson et al., 1991; Spelke and Kinzler, 2007), neurophysiology (e.g., Tsao et al., 2006;

Kriegeskorte et al., 2008), or computational modeling (e.g., Farah and McClelland, 1991; Zhuang et al., 2021)—research areas that are highly relevant to understanding the causes of occipito-temporal organization.

USE CASE SCENARIOS FOR VISUAL RECOGNITION

Consider the following example from everyday behavior: you decide to make a sandwich. You navigate to the kitchen, and on the way, all manner of objects are “perceived” but probably not noticed as such (perhaps because they are located in their typical places). In the kitchen, you get the bread and a plate and the peanut butter. You open a drawer to take out a knife. For the first time in this extended series of inner mental states and outward actions, you “visually perceive” the knife. In this example, visual perception was not the initial impetus for thinking about a “knife”; the initial proximate cause of activating conceptual, and perhaps sensorimotor, representations of “knife,” were internally represented action goals (one of which included the subgoal of using a knife to spread the peanut butter). The knife was sought (and *then* perceived) because it was *already* represented as being a useful part of a broader action goal. I refer to this as “*premeditated perception*.”

“Premeditated perception” can be contrasted with what I refer to as the “surprise paradigm.” The surprise paradigm subsumes nearly *all* experimental research on visual object recognition across a range of tasks (viewing, naming, n-back, incidental tasks), populations (human and nonhuman primates, all stages of development, healthy and lesioned), and methods (neuropsychology, functional neuroimaging, behavior, etc.). The methodological maxim of the “surprise paradigm” is that the subject should not be able to anticipate (explicitly or implicitly) what the next stimulus will be on any given trial. In that way, strategies (explicit or implicit) do not confound an interpretation of *how* the brain came to be activated the way that it did, or how the stimulus affected the participant’s behavior the way it did. The surprise paradigm, for good reason, strips away, as much as possible, the types of confounding factors that would render an interpretation ambiguous or difficult.

Within the surprise paradigm of perception, research on high-level visual processing has emphasized *identification* as the “goal” of processing in the occipital-temporal cortex. This is aligned with the posture of the system in the surprise paradigm: on each trial, a stimulus must be recognized anew, without prior context.

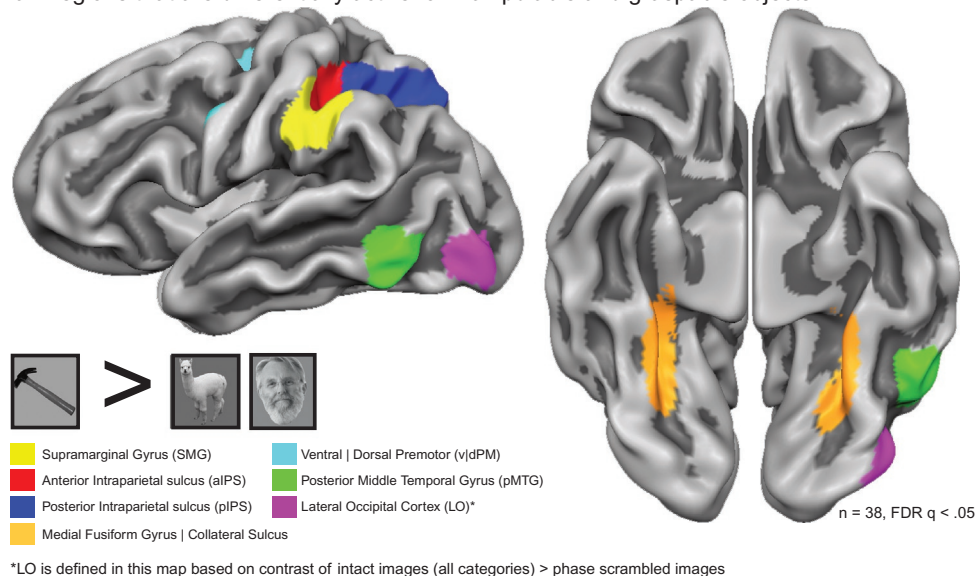
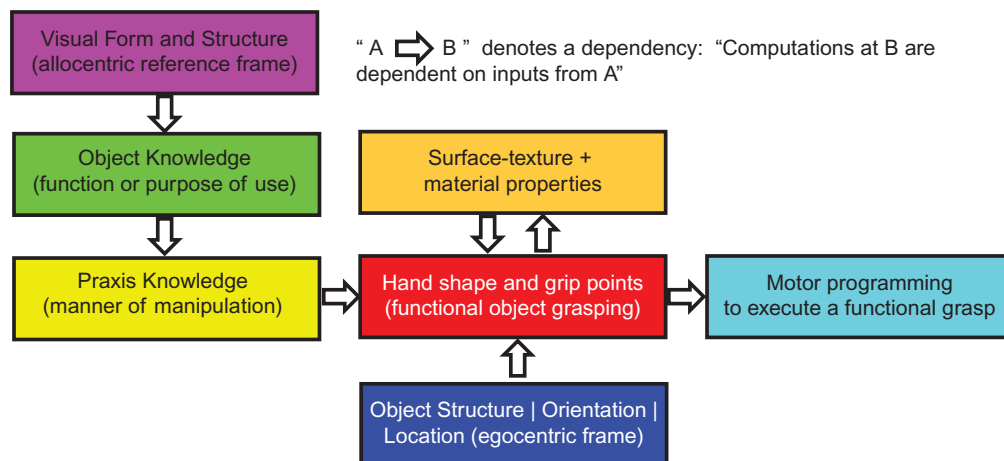
A. Brain regions that are differentially active for manipulable and graspable objects**B. Schematic of dependencies among separable computations in support of functional object use**

Fig. 13.1. (A) The maps show functional MRI data obtained while participants viewed and named images of common tools (fork, cup) compared to images of animals and faces. All of the regions, with the exception of LO, are defined as expressing differential neural (BOLD) responses for images of “tools” compared to the baseline category of “animals.” LO was defined by the contrast of all intact stimuli (tools, animals, places, face) compared to phase-scrambled images. Regions are color-coded based on the principal dissociations that have been documented in the neuropsychological literature (B). The first functional MRI studies describing this set of “tool-preferring” regions were carried out in the laboratory of Alex Martin (Chao et al., 1999a; Chao and Martin, 2000). (B) For each process/function depicted by a “box,” there are neuropsychological studies indicating that the process can be separately impaired in individuals with acquired brain injury. The schematic depicts a hypothesized series of dependencies among those dissociable processes in support of an account of “functional” object grasping (e.g., end-state comfort). The network of regions shown in (A) has been described as the “tool-processing network” (Garcea and Mahon, 2014). Calling that network the “tool-processing network” is descriptive of the types of stimuli that engage those regions. By hypothesis, the broader computational goal of the network shown in (A) is physical reasoning about how first person actions will change the state of the world (Mahon, 2020). On this analysis, functional object use is much broader than tool processing per se: “Tools,” or manipulable objects more generally, are a class of things in the world for which successful processes require the specific processes represented across the network. Reproduced from Mahon B (2020). The representation of tools in the human brain. In: D Poeppel, M Gazzaniga (Eds.), The new cognitive sciences, sixth edn. MIT Press, Cambridge, MA.

But the surprise paradigm misses an important and common “use case scenario” for visual recognition that is captured by “premeditated perception.” Objects are often already “identified” prior to their perception. We *often* do not discover what we will do with objects by looking at them; often, representations of what will be done with the object in the service of action goals are *priors* on perception (see also [Wu, 2008](#)). Consider the role of “the process of identification of the knife” in the extended sequence of making a sandwich described above. The motor-relevant information about how to use the knife to satisfy the goal of spreading peanut butter was represented independently of, and *prior* to, perception of the knife. In such cases, the first “input” to the system is not vision, but a representation of the action goal and how an object should be manipulated to accomplish that goal. “Identification,” in that context, is more a type of *confirmation* than a type of discovery.

Of course, perception is not always premeditated; perception can and does support identification, and object-associated actions *are* derivable (cold) upon visual identification of objects. In fact, pantomime of object use to visual presentation of objects, *out of context*, is a core test to assess apraxia ([Rothi et al., 1991](#)). Furthermore, the relation between motor-relevant representations of object use and high-level visual and semantic object representations is not static and fixed—it is productive and generative, and can be adapted on the fly to real-time bottom-up perceptual input ([Goldenberg and Hagmann, 1998](#)). To continue the example above, imagine that upon opening the kitchen drawer, you discover there are no clean knives; and, rather than do the dishes, you reason (*de novo*) that a spoon would do just fine for spreading peanut butter and so grasp the spoon (see [Munoz-Rubke et al., 2018](#)). This has implications when thinking about the architecture of high-level visual representations and the dynamics of how those representations interface with object use systems, on the one hand, and early perceptual processes, on the other. Whatever model of the system organization is posited, that architecture must enable two modes of physical reasoning: inferring manner of interaction and use from visual structure and general knowledge (choose the right action for a given object) and inferring the relevant visual structure in the world that enables a given behavioral goal (choose the right object for a given action).

This approach to thinking about use case scenarios for the temporal lobe takes the emphasis off of an atomization of the system into separate regions, with each region having its own function (and that function is the presumable reason for the region being a separate functional region). To separate individual regions out of the system and tell a story of selection for each region would be to

miss the relevant level of analysis for understanding the large-scale organization of the occipital-temporal cortex. Selection pressures, *if* they were relevant to understanding the nature of innate constraints that shape knowledge representations, operated on organisms, which is to say behaviors—which is to say whole networks of regions working in concert to solve computational problems. It would be a disservice to any domain-specific framework to try to fit evolutionary stories to ad hoc descriptions of functional processes in specific regions ([Gould and Lewontin, 1979](#)).

Understanding connectivity to regions outside the temporal lobe becomes the key step for understanding the organization and processing dynamics within the temporal lobe. By hypothesis, innately constrained connectivity of occipito-temporal areas to regions outside the temporal lobe is the core scaffolding that drives functional specialization by semantic domain in the temporal lobe. This approach shifts the emphasis from using neural responses to determine the semantic category tuning profile of a region, to understanding how each region contributes to the broader computational goals driving behavior. Characterizing those broad computational goals thus becomes central to the scientific enterprise of describing the functional organization of the temporal lobe.

Caution is merited with respect to “reverse engineering” the constraints that shaped the organization of the system from its current use. The current use of a system *may* be related to the original pressures that constrained the design of the system, or they may not be related. [Gould and Lewontin \(1979\)](#) explained this by analogy to spandrels—the space created between an architectural arch and the ceiling, and which is a *byproduct* of fitting a curved arch into a square space. When spandrels are adorned with a fresco, the spandrel looks like “it was designed to be decorative.” But “decoration,” while the current use, was not the original motivation for the structure. How might this apply to inferences about the causes of functional specialization in the temporal lobe?

Consider what could have been the constraints that led to the functional organization of the temporal lobe, such that there could be an area, the “visual word form area,” that exhibits specialization for printed words ([Cohen et al., 2000](#)). As will be developed further below, the visual word form area exhibits all of the hallmarks of functional specialization that a brain region might exhibit for a visual category: the visual word form area has a definite developmental timeline; focal damage causes a specific deficit (pure alexia); it has a stereotyped location across individuals, and its existence as a specialized area (for reading) is resilient to the absence of visual experience ([Buchel et al., 1998](#); [Dehaene et al., 2005](#);

Dehaene and Cohen, 2007; Striem-Amit et al., 2012; Bouhali et al., 2014; see discussion below). And yet, “reading” is the fresco that is painted on the spandrel. Reading is neither universal, nor old enough to have been a use that constrained brain organization (Dehaene and Cohen, 2007). The visual word form area is specialized for recognizing printed words, yet the visual word form area could not have been engineered for reading (Gould and Lewontin, 1979).

The moral of Gould and Lewontin’s spandrels of San Marco is clear when thinking about the visual word form area. But the visual word form area is not the exception, as much as a test-case that illustrates a methodological guardrail that should apply to *all* category-preferring areas of the ventral stream. The visual word form area is consistently in the same location across all literate individuals—presumably, in part, due to genetic constraints on the connectivity of the temporal lobe, together with other factors to do with the role of experience in shaping cortical organization (Srihasam et al., 2014). We are not inclined to assume that current use (recognizing words) was the basis for shaping whatever innate biases drive the word form area to become specialized

for reading. We should be equally cautious in assuming that current use (recognizing faces) is what drove the face area to be specialized for recognizing faces.

EVIDENCE FOR DOMAIN-SPECIFIC NEURAL ORGANIZATION

Clues from neuropsychology

Category-specific semantic deficits are impairments to conceptual knowledge that differentially, or selectively, affect information from one semantic category compared to other categories (see review by Capitani et al., 2003). Fig. 13.2 shows the picture naming performance from some well-studied patients and illustrates some of the semantic categories that can be reliably dissociated: living animate (animals), living inanimate (fruit/vegetables/plants), conspecifics (other people), manipulable objects, and not shown: geographical places and potentially body parts. The study of category-specific semantic deficits has proven to be an incredibly fertile ground for the development and evaluation of hypotheses about how conceptual information is organized in the human brain.

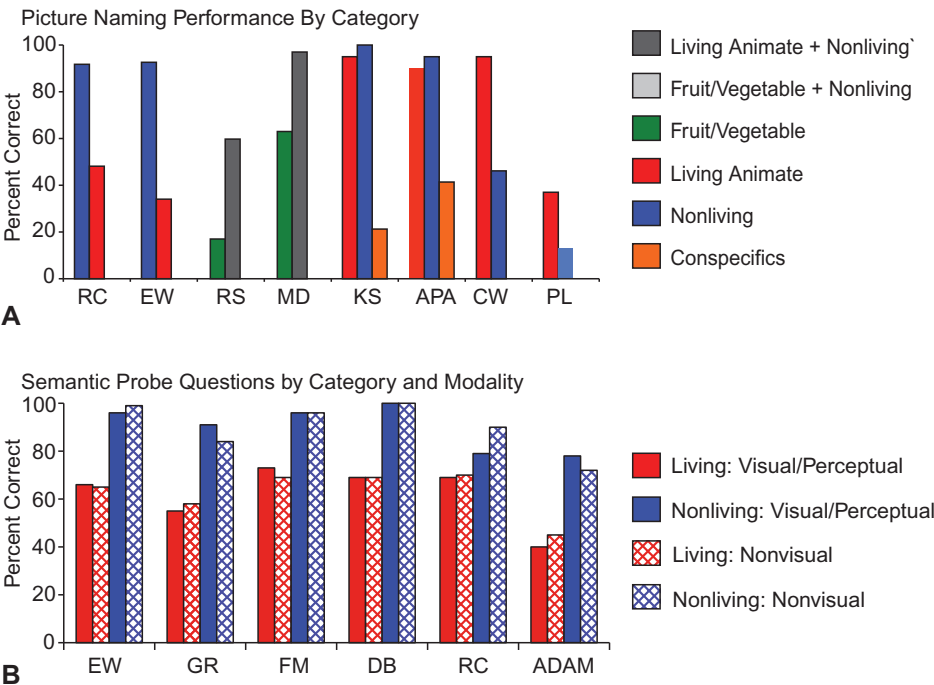


Fig. 13.2. Category-specific semantic deficits. (A) Patients with category-specific semantic deficits may have selective impairments for naming items from one category of items compared to other categories. (B) Those patients may also have impairments for answering questions about all types of object properties (i.e., visual/perceptual and functional/associative) pertaining to the impaired category (for references and discussion, see Caramazza and Mahon, 2003). Reproduced from Caramazza A, Mahon BZ (2003). The organization of conceptual knowledge: the evidence from category-specific semantic deficits. Trends Cogn Sci 7: 354–361. doi:10.1016/s1364-6613(03)00159-1.

The first important aspect of category-specific semantic deficits is that the impairment is to *conceptual* knowledge. For instance, patient EW (Caramazza and Shelton, 1998) had a selective impairment for living animate things (animals) across a range of tasks (picture naming, definition naming, sound naming, property verification). The observation that her impairment is independent of the modality of the cue (picture, sound, spoken definition) already suggests that it is not a general visual problem per se. The observation that she performs poorly on nonlinguistic visual tasks that tap into knowledge of animals but performs normally for the same tasks with nonanimal stimuli indicates that her difficulties are not reducible to either a general problem with language or vision. Category-specific semantic deficits affect *conceptual* knowledge. Of course, individual patients may show associated deficits that affect performance in one modality of input or output.

Indeed, there is evidence that category-specific semantic deficits may or may not also involve a visual agnosia for the same category of items: for instance, patient EW was impaired at determining whether the depicted chimeric animals were or were not real—so-called object reality decision. Patient KC (Blundo et al., 2006) exhibited the same functional profile of impairment, in relevant respects, as did patient EW, *except* that she performed normally in an object reality decision task involving animals and nonanimal stimuli. Thus, a category-specific deficit, selectively affecting conceptual knowledge of living animate items, can dissociate from a visual agnosia.

A similar dissociation between conceptual knowledge and visual recognition abilities has been documented with respect to knowledge and recognition of people. Patient APA (Miceli et al., 2000) had a naming impairment for people that did not generalize to all proper names, indicating it was not a general proper name anomia. APA's naming difficulties were present regardless of the modality of input (picture, voice, definition), and visual processing of faces was relatively spared. By contrast, patients with prosopagnosia are not able to visually recognize faces, especially when surrogate cues (hair, professional uniform) are not present; individuals with prosopagnosia, unlike patients such as APA, *are* able to recognize people based on nonvisual modalities of input (for instance, identifying a famous person by voice).

A second important aspect of category-specific semantic deficits is that the knowledge impairment applies to all types of knowledge about the impaired category (visual/perceptual as well as functional/associative knowledge; for a discussion, see Mahon and Caramazza, 2009). For instance, patient EW had difficulty answering questions about the visual and perceptual properties of

animals (Does a whale fly?) but no difficulty answering questions about the visual and perceptual properties of nonanimals (Is a hammer shaped like a “T”)? Similarly, EW had difficulty answering functional/associative questions about animals (Does a whale breathe air?) but no difficulty answering such questions about nonanimals (Is a hammer used by a carpenter?).

In summary, the findings from category-specific semantic deficits indicate that (i) there are a limited number of categories for which patients have been reported to have selective or disproportionate deficits at a conceptual level; (ii) category-specific semantic impairments may be accompanied (but need not be) by a corresponding category-specific perceptual agnosia; and (iii) patients with conceptual impairments have deficits for all types of knowledge (visual/perceptual, functional/associative) about the impaired category. Interestingly, and as will be discussed below, the “categories” of category-specific semantic deficits align roughly with the “categories” for which neural specificity has been demonstrated in areas of the temporal lobe, as revealed by functional neuroimaging in healthy participants.

Clues from functional neuroimaging

Over the past two and a half decades, a vibrant literature has emerged using functional magnetic resonance imaging (fMRI) to study category-specific neural organization; the early foundational studies in the late 1990s were directly motivated by category-specific semantic deficits (Chao et al., 1999b; Chao and Martin, 2000). Distilling across studies, the categories for which specific regions of the ventral visual pathway exhibit differential blood oxygen level-dependent (BOLD) responses are faces, animals, body parts, tools, places, and words (for reviews, see Grill-Spector and Malach, 2004; Martin, 2007, 2016; Op de Beeck et al., 2008; Mahon and Caramazza, 2009). On the ventral and lateral aspects of the temporal-occipital cortex, there is a consistent topography by semantic category. Viewing manipulable objects (“tools”) leads to differential BOLD contrast in the left collateral sulcus and medial fusiform gyrus, while viewing animate living things (animals and faces) leads to differential BOLD contrast in the lateral fusiform gyrus (Kanwisher et al., 1997; Chao et al., 1999b; Downing et al., 2006); for earlier work, see Allison et al. (1994). Anterior and medial, place stimuli, such as houses or scenes, as well as large nonmanipulable and highly contextualized objects (refrigerators, dressers) differentially drive BOLD responses in the parahippocampal gyrus (Epstein and Kanwisher, 1998; Barr and Aminoff, 2003) (“parahippocampal place area,” PPA). Along the medial aspect of the ventral surface of the occipito-temporal cortex, there is thus an

anterior to posterior distinction between large immovable (and nonmanipulable) objects and places driving activity in anterior regions, and manipulable objects driving activity in posterior medial ventral temporal areas (medial fusiform, collateral sulcus (Mahon et al., 2007)).

The *lateral* fusiform gyrus exhibits larger responses to face and living animate (animals) stimuli (“fusiform face area”). Interestingly, behavioral signatures of right lateralization for face processing are related to literacy: as children learn to read, face processing is biased toward the right hemisphere (Dundas et al., 2013). The area in the *left* ventral temporal cortex that comes to exhibit specialization for reading is typically in the occipital-temporal sulcus (Cohen et al., 2000), between the lateral fusiform and neighboring inferior temporal gyrus (“visual word form area”)—thus roughly homologous to face-selective areas in the right ventral temporal cortex.

There are also category effects in lateral occipital cortex for faces, body parts, hands, and objects (Downing et al., 2001; Bracci and Peelen, 2013), and subregions of those lateral occipital areas have been dissociated using TMS (Pitcher et al., 2009). More anterior, in the posterior-lateral temporal cortex, just anterior to motion sensitive area MT/V5, there is a superior-to-inferior distinction that follows an animate-to-inanimate, or a social-to-nonsocial, dimension (Beauchamp et al., 2002, 2003; Wurm et al., 2017).

The pattern of neural organization that is captured by differential neural responses to items from the above semantic categories in occipito-temporal cortex is largely invariant to the task and format of stimuli used in the experiment (e.g., linguistic, image, auditory) (Chao et al., 1999a). Category-specific responses in the ventral stream are generally resilient to stimulus transformations such as orientation, size, and contrast (Avidan et al., 2002; James et al., 2002). Specifically, the *location* of category-specific responses in occipito-temporal cortex is driven by the content of the stimulus and not its format, although the format of the stimulus and the nature of the task do modulate responses. For instance, in an early study on the topic, Chao et al. (1999a) showed that the basic living/nonliving division in the lateral/medial organization of the ventral temporal cortex was present when reading words (even though neural responses overall are more robust for images than for printed words (Chao et al., 1999a)). By contrast, the visual word form area responds to *printed* words—*any* printed word (any content). For instance, differential responses to printed words in the visual word form area are present regardless of the semantic category or content of the word, and even for pronounceable nonwords, indicating that activity in the word form area is driven by visual recognition of the printed word (for discussion, see Plaut and Behrmann, 2011).

What drives category-preferences in the ventral stream are not stimulus properties per se, but rather how the stimulus is *interpreted*. In an early and prescient study, Martin and Weisberg (2003) showed healthy participants movies of simple geometric shapes “behaving” (i.e., moving) in a manner that reflected animate things (playing chase) or inanimate objects (bouncing around like billiard balls (Heider and Simmel, 1944)). Despite the fact that the motion parameters were matched across the animate and inanimate behaviors, and that the same geometric shapes were present in both conditions, and that ventral occipito-temporal areas are not particularly motion sensitive, it was found that shapes behaving like animate things drove activity in the lateral fusiform gyrus while shapes behaving like inanimate things drove activity in the medial fusiform area. That shows it is not the shapes or even their motion trajectories that drive bottom-up activity in category-preferring regions—rather, it is the interpretation of what the shapes were *doing*. Such interpretations do not (by hypothesis) depend on processing in a single brain region—those interpretations depend on inferences being drawn based on processing across many brain regions (see below and Caramazza and Mahon, 2003; Martin, 2016).

There is thus a consistent pattern of organization across the ventral and lateral surfaces of the occipito-temporal cortex, defined by regions that exhibit a peak, and roughly selective responses to the categories of faces, animals, places, tools, printed words, body parts, and potentially numbers. Some regions are known by the category of stimuli that maximally drive responses, for instance, the fusiform face area, the parahippocampal place area, and the visual word form area. The names pick out the tuning profiles of the regions. Knowing the tuning profile of the region is important fundamental knowledge. And, having criteria for functionally defining a brain region in an agreed upon manner allows different researchers to be confident they are studying the same functional part of the part of the brain across individuals, and across time within individual participants followed longitudinally (see discussion by Friston et al., 2006; Saxe et al., 2006). But the naming conventions are descriptive, not explanatory. As such, the core substantive questions are only just framed: What are the [phylogenetic | ontogenetic | real-time processing] constraints that give rise to a consistent anatomical distribution of category-preferring regions across individuals?

Perceiving and knowing about color and surface texture

The discussion to this point has emphasized a way of thinking about vision as primarily informing “what” a

visual stimulus is. That approach naturally highlights the role of visual form and shape analysis in support of visual categorization, identification, and naming. As [Conway \(2018\)](#) (see also [Carroll and Conway, 2021](#)) puts it, while visual form tells us *what* something is, color and surface texture tell us why we care about it. The color and surface texture of a face, fruit, or an object inform its relevance to behavioral goals. If the goal is to infer something about what someone is feeling, perceiving the blush in their face is highly relevant ([Changizi et al., 2006](#)); see [Carroll and Conway \(2021\)](#). By contrast, for instance, surface texture cues that support inferences about slipperiness or weight would be highly relevant to the goal of functional object grasping.

Color is not an attribute that is detached from the rest of vision, although it is detachable: under conditions of genetic abnormality, acquired injury, or certain psychophysical conditions, color processing and knowledge can be dissociated from both form and motion ([Livingstone and Hubel, 1988](#); [Siuda-Krzywicka and Bartolomeo, 2020](#); [Carroll and Conway, 2021](#)). That color perception and knowledge are functionally dissociable from other types of visual processing provides clues about how different parallel pathways in the visual system integrate processing. Color perception, declarative knowledge of the typical colors of typically colored things, and the ability to name colors are supported by at least partially dissociable systems ([Miceli et al., 2001](#); [Stasenko et al., 2014](#); [Siuda-Krzywicka et al., 2019, 2020](#)); see reviews in [Carroll and Conway \(2021\)](#) and [Siuda-Krzywicka and Bartolomeo \(2020\)](#). Lesions that involve the ventral medial occipito-temporal areas are associated with achromatopsia when they involve more posterior segments (occipital lingual gyrus), and color agnosia when they involve more anterior aspects ([Miceli et al., 2001](#); [Stasenko et al., 2014](#); [Siuda-Krzywicka et al., 2019](#)) such as the anterior lingual gyrus, collateral sulcus, and medial fusiform gyrus. Patients with achromatopsia but not color agnosia have difficulty with color perception but do not necessarily have difficulty with knowledge of the typical colors of things (e.g., knowing that grass is green, or that a watermelon is a different color on the inside than on the outside). By contrast, patients with color agnosia without achromatopsia can be spared for color perception but impaired for color knowledge. And importantly, both color agnosia and achromatopsia are observed in the setting of spared visual form processing. (For a discussion of issues related to color knowledge, see also [Chapter 10](#), this volume.)

Functional MRI studies have confirmed the major neuropsychological dissociations (e.g., [Cant and Goodale, 2007](#); [Simmons et al., 2007](#)). [Simmons et al. \(2007\)](#) found that the region of the ventral occipito-temporal cortex (along the lingual gyrus) that was engaged when retrieving

knowledge of object color is just anterior to the region involved in color perception. [Siuda-Krzywicka et al. \(2021\)](#) used resting functional connectivity in healthy participants to study the behavioral relevance of dissociable functional networks associated with independently localized posterior and anterior color regions in the ventral temporal cortex. Those authors found that network connectivity of posterior color regions predicted response time variance in color categorization, while network connectivity of anterior color areas predicted response time variance in color naming—again in broad agreement with the neuropsychological data.

The above-described neuropsychological studies found that color and form processing dissociate; of course, some occipito-temporal lesions can also cause associated deficits for color and form. Certainly, at the extreme, lesions involving the primary visual cortex eliminate all experience of seeing both form and color—although residual “perceptual” and action abilities can be demonstrated in the hemianopic field of such patients based on rudimentary form processing ([Prentiss et al., 2018](#)), such performance is without phenomenal awareness of a definable visual percept. An exception, interestingly, is that some patients can still experience motion in their hemianopic field (the “Riddoch” phenomenon ([Zeki and Ffytche, 1998](#)); see also [Stoerig and Cowey \(1989\)](#) for tests of spectral sensitivity in the hemianopic field).

In the normal course of typical experience, color and surface texture are bound up with our experience of other visual dimensions. This is not in conflict with demonstrations that color perception and color knowledge, and visual form perception and form knowledge, are mutually dissociable. At sufficiently early stages of processing, color and form are processed in retinotopic coordinates, and retinotopy thus provides the fabric for the integration of color and other surface properties with form (see discussion by [Conway, 2018](#); [Carroll and Conway, 2021](#)). At subsequent stages of processing, object representations are no longer in retinotopic frames—and color *can* become separated from its object ([Friedman-Hill et al., 1995](#)).

Importantly, color is represented at multiple points along occipito-temporal pathways, concentrated along the lingual gyrus, collateral sulcus, and medial fusiform gyrus, between place responsive areas medially and face responsive areas laterally ([Lafer-Sousa and Conway, 2013](#); [Lafer-Sousa et al., 2016](#)). As will be discussed below, it is of note that the location of those color sensitive areas seems to roughly align with the location of regions that exhibit neural specificity for manipulable objects ([Chao et al., 1999a](#)). I will argue that biases to represent manipulable objects in those specific regions (collateral sulcus, medial fusiform gyrus) reflect the need

to integrate representations of objects' surface texture with object-directed action representations processed in parietal regions (Mahon, 2020).

For faces and perhaps edible plants, color is a highly relevant cue that supports the broader computational goals demanded of behaviorally relevant processing in those domains—it supports the “why we care (Conway, 2018).” For manipulable objects that are grasped and used for a purpose, color is (typically) an irrelevant dimension. A hammer is a hammer regardless of its color. That is to say, the types of inferences associated with grasping and using a hammer are not different depending on the hammer's color—as long as color is not a cue to other properties that would affect how one could interact with the object. In particular, surface texture is a cue to infer the coefficient of friction at the object's surface (grip force) and the material composition and weight density of the object (grasp location). Lesions involving the collateral sulcus have been found to disrupt processing of surface texture, without affect processing of visual form (Cavina-Pratesi et al., 2010; see Cant and Goodale, 2007 for convergent evidence from fMRI). Furthermore, inferences about material composition and object weight have also been shown (Gallivan et al., 2014) to be supported by the collateral sulcus and adjacent structures (lingual gyrus, medial fusiform gyrus).

EXPLANATIONS OF CATEGORY-SPECIFICITY IN OCCIPITO-TEMPORAL CORTEX

Local-constraint-based accounts of the causes of category-preferences

The power of using functional neuroimaging to study category-preferences is that it provides a window to *all* regions that are involved in processing information about a category, regardless of whether involvement of those regions is necessary. This offers a useful complement to neuropsychological evidence, which emphasizes the contribution of processes that are *needed* for intact performance. There are many accounts that have been proposed to explain why category-preferences in the ventral stream have a consistent anatomical organization (for instance, Gauthier et al., 1999; Grill-Spector and Malach, 2004; Rogers et al., 2005; Mechelli et al., 2006; Martin, 2007, 2016; Op de Beeck et al., 2008; Mahon and Caramazza, 2011; Plaut and Behrmann, 2011; Lafer-Sousa and Conway, 2013; Conway, 2018).

Most research on category-preferences in the ventral stream has worked under the empirical precept that the category for which a given subregion exhibits specificity is just the category of items that elicits a disproportionate

(or selective) neural response (compared to items from other “categories”) (Downing et al., 2006). The assumption that response *amplitude* is a necessary and sufficient criterion for inferring regional specialization derives from traditions in neurophysiology and is in line with a broader paradigm for thinking about the causes of ventral stream organization (Downing et al., 2006).

One class of proposals, local-constraint-based proposals, makes two broad assumptions. First, the processing in the ventral stream that gives rise to category-preferences is a stimulus-driven analysis of the sensory input. Such a “stimulus-driven” analysis may involve feedback—that it is “stimulus driven” does not mean that it is “feedforward only” in terms of processing (Riesenhuber and Poggio, 1999). Second, local-constraint theories emphasize that category-specificity is due to constraints that are expressed only over information represented local to the brain area showing category-specificity, or perhaps *earlier* in the processing hierarchy than the stage at which the category-preferences manifest. A more middle-ground version of the local-constraint view may emphasize connectivity among occipito-temporal areas more generally—for instance, giving rise to competition between homologous regions in the left and right hemispheres for reading and face recognition, respectively (Plaut and Behrmann, 2011).

The most widely endorsed local-constraint approach for thinking about the causes of category-preferences in the ventral stream is that category-specificity emerges from the interaction of statistical regularities in visual experience with domain-general principles of organization. Some proposals emphasize dimensions that have been shown to account for some, but not all, category effects. For instance, Rogers et al. (2005) argued that processing of animals requires more fine-grained processing than tools; that theory (unadorned) stumbles to account for the double dissociation (i.e., regions that exhibit differential responses to tools and large nonmanipulable objects than animals). As another example, Mechelli et al. (2006) argued that the “relevance” of semantic features is unevenly distributed across categories, and that relevance is correlated with BOLD signal; this proposal has difficulty with double- (and triple-order) dissociations. More recently, object size has been proposed as a constraint on organization in ventral temporal cortex organization (Konkle and Oliva, 2012). An ambiguity of that proposal is that it is not clear if it is size, *per se*, or other dimensions that align with size (e.g., motor-relevant dimensions, navigational relevance).

Another group of proposals argues that categories differ on mid-level visual features, such as having curvy edges or straight edges, and that the “original” bias

(“proto-maps”) in the ventral stream is by such mid-level features, with category-preferences being secondary to that more basic organization by feature type (Nasr et al., 2014; Arcaro and Livingstone, 2017). Along those lines, perhaps the most developed and empirically supported proposal is that category-preferences arise because higher-order visual areas inherit weak retinotopic biases from earlier visual areas—faces are represented in the part of higher-order visual cortex that receives differential input from para-foveal regions, while “places” are represented in the part of higher-order visual cortex that receives differential input from peripheral visual field locations (and, by hypothesis, we tend to fixate faces) (Levy et al., 2001; Hasson et al., 2002; Conway, 2018). Finally, “high-dimensional” theories have been sketched, where multiple local dimensions might combine in possibly nonlinear ways to yield the observed patterns of category-preferences (Op de Beeck et al., 2008).

I refer to the above proposals, in cohort, as “*local-constraint* theories” to emphasize the common feature across those theories: dimensions of organization *local* to the ventral stream interact with regularities in the visual input to give rise to category-specificity. There is evidence for each of the local constraint theories briefly reviewed above—the argument here is not that retinotopic preferences or mid-level shape primitives do not map onto the organization of occipito-temporal areas by semantic category; nor is the argument that such dimensions are not causally relevant in explaining the development and tuning of the system. The argument is that such constraints are not enough—they do not account for all of the relevant evidence. What is missing on such accounts is connectivity between occipito-temporal areas and regions outside the temporal lobe.

Connectivity-based explanations of the causes of category-preferences

An alternative to local-constraint-based theories is the proposal that the large-scale organization of the ventral stream by domain is a consequence of its (innately constrained) connectivity with other brain regions. We refer to this class of proposals as “connectivity-constrained accounts,” following the term introduced by Riesenhuber (2007). On this view, category-specificity is a consequence of occipito-temporal regions being innately constrained to be connected to other brain regions in the service of computational goals. If there were selection pressures that shaped how the brain processes and represents the world, those selection processes did not operate over regions, or the processes supported by individual regions. They operated over

behaviors that in turn depend on the integrated functioning of many regions, in networks, that are oriented toward solving different types of computational goals. As a proposal about the constraints that shape neural organization, a connectivity-based account is complementary to local-constraint-based approaches. That said, there are findings that fit more naturally within a connectivity-constrained account, and there are findings that are simply incompatible with a local-constraint-based account.

An interesting possible *extension* on the connectivity-constrained account of the organization of the ventral stream is the idea that real-time processing in the ventral stream may causally depend on the real-time computations carried out by regions outside the ventral stream (Price et al., 2001; Garcea et al., 2019). This ties into the origins of the connectivity-constrained approach, at least as applied to occipito-temporal organization around categories. Using repetition priming in fMRI to probe neural specificity, we found that properties of manipulable objects that have to do with the way in which they are physically manipulated modulate the patterns of neural responses in specific regions of the ventral stream (Mahon et al., 2007; see also Valyear and Culham, 2010). Specifically, stimulus-specific repetition suppression (a form of neural repetition priming) in the medial fusiform gyrus was stronger for manipulable objects with a direct mapping between the manner of manipulation and visual structure and function (i.e., tools) compared to objects for which there was an arbitrary or variable mapping (arbitrarily manipulated objects). Those data can be accounted for by assuming that neural specialization for manipulable objects in the ventral occipito-temporal cortex is driven in part by real-time interactions between that region of the ventral stream and parietal areas that support object-directed action (Valyear et al., 2012; for a computational demonstration, see Chen and Rogers, 2015). More broadly, a connectivity-constrained account has been extended to other classes of stimuli that drive category-preferring responses, including faces, words, and body parts (Bracci et al., 2010, 2012; Saygin et al., 2011; Bracci and Peelen, 2013; Bouhali et al., 2014; Osher et al., 2016; Saygin et al., 2016; Li et al., 2020). For instance, it has been shown (Martin, 2006; Bouhali et al., 2014) that connectivity of the visual word form area with left hemisphere language centers is related to category-preferences for printed words (for an early suggestion along these lines, see Martin, 2006).

Stepping back, a core commitment of a connectivity-constrained account is that the constraints that drive specialization of function are not limited to being expressed over “visual” information. This suggests that, in order to understand regional specialization in

occipito-temporal areas, it is necessary to understand which regions *outside* the ventral stream are connected to category-preferring regions within the ventral stream, and the types of computational problems that are solved by those networks. This approach shifts the emphasis from understanding which stimuli maximally drive responses in a region, to the connectivity profiles of ventral stream regions and how those connectivity profiles align with functional response profiles. In short, each ventral stream region (/voxel) will have a “semantic category tuning curve” and a “connectivity tuning curve.” A connectivity-constrained account predicts that they should be in alignment; and the proposal that connectivity is innately constrained predicts that alignment should be resilient to an absence of visual experience.

ADJUDICATING BETWEEN LOCAL-CONSTRAINT- AND CONNECTIVITY-BASED ACCOUNTS OF CATEGORY-SPECIFIC NEURAL ORGANIZATION IN THE VENTRAL STREAM

Category-preferences are present in congenitally blind individuals

If the constraints that shape functional specialization in the ventral stream are based on connectivity with regions outside the ventral stream, and that connectivity is innately constrained, then category-specific neural organization in the ventral stream should be present even in the complete absence of visual experience. By contrast, *local*-constraint-based accounts are committed to the view that the organization of the ventral stream will be determined by how *visual* information is processed in occipito-temporal areas. If the same patterns of category-preferences are observed in brains that have never had visual input, then we can conclude that visual experience is not necessary to produce that organization. Moreover, confidence would increase for the inference that what supports the emergence of category biases in the ventral stream is hardwired connectivity.

Four findings have stood the test of replication: First, large nonmanipulable objects, which may be understood as being “navigationally relevant” landmarks (Barr and Aminoff, 2003), differentially activate anterior medial regions of the ventral temporal cortex both in sighted and congenitally blind individuals (Mahon et al., 2009; He et al., 2013). Second, reading, whether with the fingers (Braille reading in the congenitally blind) or the eyes (visual reading of printed words in sighted individuals), activates the same region of the left lateral ventral

occipito-temporal cortex—along the occipito-temporal sulcus (Buckley et al., 1998; Striem-Amit et al., 2012). Third, manipulable objects (“tools”) activate parietal action-related areas with the same anatomical distributions in sighted and blind brains (Mahon et al., 2010; Peelen et al., 2013). Interestingly, and following the pattern seen in sighted participants, action-associated parietal activity was left lateralized in inferior parietal areas. Fourth, in the lateral occipito-temporal cortex, the left posterior middle temporal gyrus (aka LOTC) is driven by action-relevant conceptual processing, both in sighted and in blind subjects (Kemmerer and Gonzalez-Castillo, 2010; Peelen et al., 2013). Those four findings indicate that the organization of the respective systems depends on constraints that are independent of (or resilient to variance in) visual experience: the constraints that are *sufficient* to drive category-preferences are *endogenous* to the brain, independent of visual experience.

It is important to underline what is not being argued: A connectivity-constrained account is not arguing that “experience does not matter”; it is also not denying the visual nature of the information represented in visual areas in sighted individuals. As an example, consider patient DF, who had bilateral lesions to lateral occipital cortex (LO), and a dense visual form agnosia—those data, followed by a wave of fMRI studies, have collectively shown that LO represents visual shape. In congenitally blind individuals, LO has no visual input, and yet it still represents shape, as assessed through touch and hearing using sensory-substitution devices (Amedi et al., 2001, 2007). Thus, LO represents shape, even via nonvisual modalities when vision is not an input to the system. Similarly, the visual word form area represented printed words, presumably in a tactile format for Braille readers.

What is being argued: the *locations* of category-preferences do not depend in any strong (or interesting) sense on visual experience, and which categories are represented in those locations does not depend on visual experience. Of course, the format of information that is represented in those occipito-temporal regions is *determined* by experience (visual information, tactile information). What is the “same” between sighted and blind brains is not the information that is represented but the anatomical location in which that *type* of information is represented. By inference, this is because the computations demanded of that *type* of stimulus are the same across sighted and blind individuals. Those computations are supported by innately constrained connectivity between subregions of the ventral stream and other regions of the brain (Mahon et al., 2007, 2009; Mahon and Caramazza, 2011).

Alignment of connectivity with category-preferences

A premise of this review is that clues about what ventral occipito-temporal areas are *doing*, and why those regions have the processing preferences that they do, are provided by understanding the connectivity of those occipito-temporal areas to the other brain regions. It is not enough to just test which category of stimuli elicits the maximal response in a region—although that can be a key starting point for motivating or “seeding” a subsequent analysis of connectivity. Understanding the connectivity of each occipito-temporal region provides clues as to *why* it has the semantic-category “tuning profile” that it does.

Taking this a step further, a connectivity-constrained account should prioritize connectivity as a criterion for determining whether a given region is part of a domain-specific network. In other words, the value of studying connectivity is not just to understand how certain regions (defined by stimulus preferences) are connected. That approach still considers maximal neural responses as the defining criterion for functionally parcellating occipito-temporal areas. Parenthetically, it is valuable to recognize that the structure of our current theories is likely partly due to the contingency of the technology we have had as a field to study the brain, and the order in which that technology has come into widespread use. The widespread use of functional MRI to discover functional areas, and the widespread use of univariate contrasts, positioned that type of data as the starting point for investigations into connectivity. Those investigations into connectivity are now coming into their own and providing new and exciting insights. But it could have been the other way around. It could have been that we started with a detailed understanding of the connectivity of occipito-temporal areas and *then* started to test the specificity of cortical organization. Imagine how different would be our theoretical frameworks and reasoning processes if we were seeking to understand how known patterns of connectivity mapped onto (unknown) stimulus preferences. That is akin to the mental flip for which I am advocating.

The most basic expectation of a connectivity-constrained account is that there will be privileged structural and functional connectivity between category-preferring regions in the ventral stream and regions outside of the ventral stream that process information about items from those categories. For instance, in a particularly elegant study, Bouhali et al. (2014) found that white matter connectivity of the visual word form area to known language areas in the left hemisphere predicted (and, by inference, *constrained*) the location in

occipito-temporal cortex where functional preferences for printed words are observed.

A number of studies have also now documented alignment between “category-preferences” and connectivity (Mahon et al., 2007, 2013; Hutchison et al., 2012; Gallivan et al., 2013a,b; Stevens et al., 2015; Garcea et al., 2018). Recent findings support a connectivity-constrained account by demonstrating *voxel-wise* alignment between patterns of connectivity (structural and functional) and locations of category-specific functional preferences. In other words, while prior demonstrations showed that *regions* that exhibit similar category-preferences are functionally connected, the findings to be reviewed next showed that the distribution of category-preferences over ventral stream voxels is systematically related to the distribution of connectivity values over the same voxels.

In one demonstration of such alignment between functional response characteristics and connectivity (Chen et al., 2017), a group of healthy participants were scanned while they viewed tools, animals, faces, and places (Fig. 13.3). Resting fMRI scans were obtained for each participant. Regions of interest (ROIs) were defined in (i) parietal areas that preferred tools to non-tools (tools > faces and places), (ii) the right superior temporal sulcus area that preferred faces to nonfaces (faces > tools and places), and (iii) the retrosplenial cortex that preferred places to nonplace stimuli (places > faces, tools, and animals). Thus, we first defined the regions *outside of the ventral* stream with which, by hypothesis, category-preferring regions in the ventral stream should be connected. For each subject, resting functional MRI data were used to compute whole-brain functional connectivity maps using the parietal (tool), STS (face), and retrosplenial (place) ROIs as seeds. We then correlated, across ventral stream voxels, the multivoxel pattern of functional connectivity to parietal, STS, and retrosplenial cortex areas with tool, face, and place preferences. We found that in medial regions of the ventral stream, the multivoxel pattern of tool preferences was positively correlated with the multivoxel pattern of functional connectivity to the parietal cortex, but not with the multivoxel patterns of functional connectivity to the retrosplenial cortex or STS. Similarly, place preferences in the medial ventral stream were positively correlated with the multivoxel pattern of functional connectivity to the retrosplenial cortex, but not with the multivoxel pattern of connectivity to the parietal cortex or STS. And finally, for the third component of the triple order dissociation: face preferences in lateral regions of the ventral stream were correlated with the multivoxel pattern of connectivity to STS, but not with the multivoxel pattern of functional connectivity to the

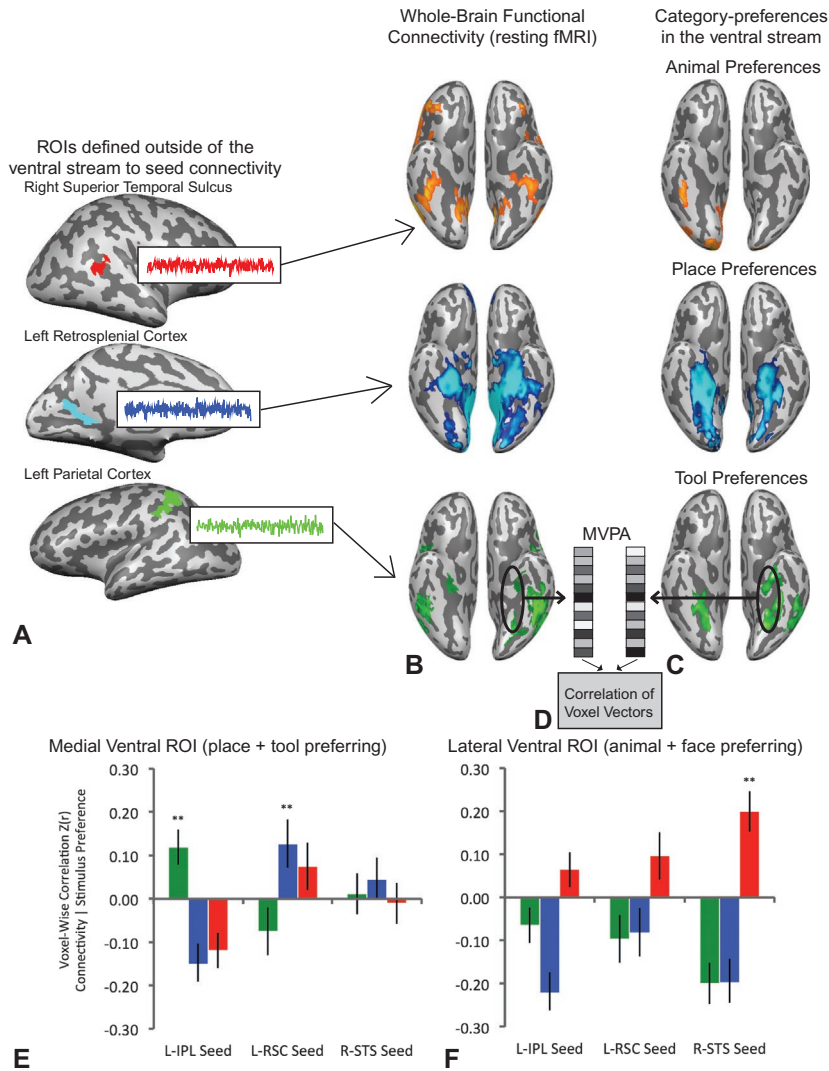


Fig. 13.3. Alignment of category-preferences in the ventral stream with functional connectivity to category-preferring regions outside of the ventral stream. (A) Category-preferring regions outside ventral occipito-temporal cortex were defined for animals (superior temporal sulcus), places (retrosplenial cortex), and tools (parietal cortex). (B) Whole-brain functional connectivity maps were generated over resting fMRI data using each of the Regions of Interest (ROIs) outside of ventral occipito-temporal cortex as seeds (the figure shows the results only for ventral occipito-temporal cortex). (C) Separately, category-preferences for animals, places, and tools were computed in the ventral stream. (D) Linear correlation was used over ventral stream ROIs to relate the multivoxel pattern of functional connectivity (to regions outside the ventral stream) to category-preferences. (E) In medial ventral ROIs (place and tool-preferring areas), the multivoxel pattern of functional connectivity to retrosplenial cortex was related to place preferences but not tool preferences, shown in the *blue bars*; over the same pool of medial ventral stream voxels, the multivoxel pattern of functional connectivity to parietal cortex was related to tool preferences but not place preferences, shown in the *green bars*. (F) In lateral ventral ROIs, the multivoxel pattern of functional connectivity to the superior temporal sulcus was related to animal (and face, not shown) preferences (*red bars*), but not to tool or place preferences.

parietal cortex or to the retrosplenial cortex. What those findings show is that the multivoxel distribution of resting functional connectivity to regions outside the ventral stream that process faces, tools, and places is related to the multivoxel pattern of stimulus preferences within the ventral stream.

Zooming out, another line of studies took a more agnostic approach as to *which* regions/voxels outside

of the occipito-temporal cortex exhibit differential connectivity to known category-preferring regions in occipito-temporal cortex. Osher et al. (2016) demonstrated voxel-wise alignment between structural connectivity and category-preferences, and Saygin et al. (2016) found that patterns of structural connectivity presage functional preferences for faces and printed words (see also Li et al., 2020). Both of those studies

computed the whole-brain pattern of connectivity for each occipito-temporal cortex voxel, rather than testing specific hypotheses about which regions outside the ventral stream expressed privileged connectivity to ventral stream areas.

Macroscopic alignment in functional connectivity between category-preferring regions in the ventral stream and regions outside of the ventral stream is, of course, not in itself incompatible with local-constraint-based theories of the causes of ventral stream organization. For instance, it could be argued that the organization of the ventral stream by category depends on statistical regularities of mid-level visual features across categories (Nasr et al., 2014), such that the connectivity of the system emerges *secondary* to the constraints that drive the organization of the system (Conway, 2018). But to flip this around, it is also the case that accounts based on mid-level visual features do not *predict* the existence of such connectivity patterns, whereas a connectivity-constrained account is committed to such predictions. Thus, taken in the context of other findings, for instance, the data from congenitally blind subjects, the existence of privileged functional connectivity between category-specific regions in the ventral stream and regions outside the ventral stream is more in line with what would be expected based on a connectivity-constrained account, compared to a local-constraint account.

The studies just reviewed are constraining in terms of the anatomy of alignment of connectivity and category-preferences—but they do not directly address the key question of which came first: does connectivity constrain category-preferences or does the connectivity emerge secondary to category-preferences (with the category-preferences established via local constraints). One recent study offers a suggestion along those lines: Kamps et al. (2020) found that functional connectivity was present in neonates within networks of regions that would later go on to develop face and place specificity—indicating that signatures of domain-specific connectivity are present already at birth. This is clearly a space in which developmental cognitive neuroscience data will play an outsized role in adjudicating which aspects of organization are explained by a connectivity-based account, and which are explained by local-constraint-based accounts.

Is the connectivity that drives category-specific organization innately constrained?

The studies just reviewed are in line with the prediction made by a connectivity-constrained account that it is the connectivity of ventral stream regions with regions outside of the ventral stream that provides the initial

scaffolding for category-specificity in the ventral stream. The other core component of this proposal, within a domain-specific framework, is that that connectivity is *innately* constrained. It is important to separate, at an evidentiary level, the hypotheses that (i) category-preferences depend on connectivity constraints, and (ii) connectivity constraints that lead to category-preferences are innately constrained.

What is the evidence that the connectivity constraints that shape category-specificity in occipito-temporal regions are innately constrained? Some of that evidence is circumstantial and indirect—but reviewed here together, there are five convergent lines of evidence that increase confidence in connectivity constraints being innate.

First, the “typical” macroscopic pattern of category-specific organization of occipito-temporal areas is present in congenitally blind subjects. That indicates, minimally, visual experience is not necessary for that organization to be present. What else, but something to do with connectivity, could explain the resilience of category-preferences to variation in sensory input (e.g., vision to tactile)?

Second, in typically reared nonhuman primates, there are well-circumscribed face patches (Tsao et al., 2006). Recent rearing studies that have prevented developing macaques from seeing human faces have shown that in those situations, face specificity fails to develop (Arcaro et al., 2017). Those data indicate that experience is relevant (and yet the data from the congenitally blind suggest it is not necessary). The inference to be extrapolated is that if the system is not exposed to the relevant experiences at critical periods in development, it will fail to develop. Showing that experience is relevant at key stages in development indicates innately constrained learning processes that are unfolding (and which can be derailed).

Third, studies such as that of Kamps and colleagues suggest that even in neonates, there are patterns of connectivity that anticipate what will become category-preferring areas (see also Saygin et al. (2016) for a demonstration that connectivity predicts the *future* location of the visual word form area). More broadly, there is a long-established developmental argument that infants come to their experiences with a repertoire of innately constrained learning systems that are specialized for acquiring and representing knowledge about things and events from different domains (e.g., agents, physical reasoning about the world; Spelke and Kinzler, 2007).

Fourth, category-specific deficits can be present from birth, and surprisingly, remarkably resilient to recovery, even in the context of an otherwise functional visual system. For instance, a category-specific semantic deficit

for living things was observed to arise from a brain injury that occurred at 1 day of age (with the patient tested at age 16; Farah and Rabinowitz, 2003). With respect to face processing, there is a robust literature documenting “lifelong” or “congenital” prosopagnosia (Duchaine and Nakayama, 2006); a recent and exhaustive review of the literature (Geskin and Behrmann, 2018) found that 80% of reported cases had a concomitant impairment for nonface categories (e.g., reading printed words) while 20% seemed to have (at least based on current evidence) a selective impairment restricted to face processing. Prior evidence (Thomas et al., 2009) indicates that individuals with congenital prosopagnosia have disrupted structural connectivity between posterior and anterior temporal lobe regions. Within this context, a finding that requires explanation is the observation that individuals with congenital prosopagnosia can exhibit seemingly normal neural responses to faces in posterior ventral stream areas (e.g., Hasson et al., 2003).

Fifth, there is greater similarity in category-specific neural organization between monozygotic than dizygotic twins. Polk et al. (2007), and subsequently a larger study by Abbasi et al. (2020), found that neural signatures of category-specificity are more similar between monozygotic twin pairs than dizygotic twins. At a behavioral level, Wilmer et al. (2010) (see also Zhu et al., 2010) found greater similarity in face processing abilities for monozygotic twin pairs than for dizygotic twin pairs. Those twin studies indicate a genetic contribution to neural and behavioral phenotypes of category-specificity—outcomes that are predicted by the proposal of innately constrained connectivity-based constraints. Another expectation along these lines is that sulcus and gyral folding patterns and their relation to white-matter connectivity should be systematically related to some category-preferences (Natu et al., 2021).

Causal evidence for a connectivity-constrained account of category-preferences

In the course of everyday action, object grasps are semantically informed—they are calibrated to *what* it is that is being grasped, and to the surface texture and material composition of the object—which is to say, how it should be it grasped. In order to grasp an object in a functional manner, by the appropriate part of the object and with appropriate force, information about surface-texture and the material composition and weight distribution of the object must be taken into account by the processes that determine hand and finger posture and grip strength. Anterior IPS (aIPS, Fig. 13.1) supports

hand shaping in the service of object-directed grasping (Binkofski et al., 1998; Culham et al., 2003; Mruczek et al., 2013). The medial fusiform and adjacent collateral sulcus and lingual gyrus (i) support analysis of surface textural properties and inferences about material composition (Miceli et al., 2001; Cant and Goodale, 2007; Simmons et al., 2007; Cavina-Pratesi et al., 2010; Gallivan et al., 2014; Stassen et al., 2014) and (ii) exhibit differential neural responses to “tools” compared to animals and faces (Chao et al., 1999a; Mahon et al., 2007).

The dorsal visual pathway processes visual inputs in parallel to analysis within occipito-temporal pathways, and receives both subcortical projections that bypass the primary visual cortex as well as inputs via the primary visual cortex. The dorsal visual pathway, on its own, can compute semantically uninformed object grasps that respect the volumetric properties of the object and its real-world location in body- or retinotopic coordinates. In other words, the dorsal stream, on its own, does not have access to information about *what* the object is, or to information about the material composition and weight distribution of the object.

By hypothesis (Mahon, 2020), category-preferences for tools in the medial fusiform gyrus and collateral sulcus are a reflection of the interactions that allow the system to direct the correct actions to the correct parts of the correct objects. Specifically, category-preferences for “tools” in the medial fusiform gyrus and collateral sulcus result from two intersecting constraints: access to surface-texture and material properties via a bottom-up analysis through the ventral visual hierarchy, and inputs from dorsal stream regions that are computing grasp-relevant parameters and which thus need to “know” about those object properties (see Mahon, 2020). There is no reason to believe that inputs from aIPS to medial ventral stream regions are “top-down”—that pathway is (by hypothesis) an aspect of how the system initially processes visual information in the service of object-directed action (see Bar et al., 2006).

The proposal that neural responses to “tools” in medial ventral stream regions are the result of joint inputs from the ventral visual hierarchy and the dorsal visual pathway predicts privileged connectivity between the medial fusiform gyrus and aIPS—a pattern that is now well attested across studies and labs (Mahon et al., 2007; Gallivan et al., 2013b; Garcea and Mahon, 2014; Stevens et al., 2015; Chen et al., 2017; Garcea et al., 2018). That proposal also predicts that stimulus factors that modulate activity in parietal action areas should have echoes on neural activity in the medial fusiform gyrus and collateral sulcus. That prediction is in line with observations that action-relevant properties

of objects modulate neural responses in the fusiform gyrus (Mahon et al., 2007; Chen et al., 2018). But perhaps one of the strongest predictions that can be generated is that lesions to aIPS should have a direct modulatory effect on neural responses in the medial fusiform gyrus and collateral sulcus. Specifically, lesions to aIPS should *reduce* responses in the medial fusiform/collateral sulcus to “manipulable object” stimuli but should not affect responses to other classes of stimuli in the same region (faces, animals, and critically—places). In other words, evidence for specificity is provided by studying the functional consequences of disrupting processing in brain regions outside the occipito-temporal cortex.

Garcea et al. (2019) tested the prediction that parietal lesions will affect neural responses (specifically) to manipulable objects in (specifically) the medial fusiform/collateral sulcus, using a novel technique termed “Voxel-Based Lesion Activity Mapping (VLAM). VLAM is similar to the well-established approach of Voxel-Based Lesion Symptom Mapping (VBLSM; Bates et al., 2003). In VBLSM, each patient from a large group contributes a lesion mask and a performance measure on a neuropsychological task of interest. The likelihood of any voxel being lesioned across the group is correlated with (regressed on) performance on the neuropsychological task. VBLSM thus provides a map of where the presence or absence of lesions predicts variance in the reference neuropsychological task. By contrast, in VLAM analyses, each patient contributes a lesion mask and a whole-brain map of stimulus-evoked activity, for instance, a whole-brain map of neural responses to “manipulable objects.” The core finding from Garcea and colleagues was that the amplitude of tool responses in the medial fusiform gyrus and collateral sulcus was inversely related to the likelihood of lesions involving aIPS (Fig. 13.4). By contrast, the amplitude of place preferences in the same collateral sulcus and medial fusiform region was not related to presence of lesion in aIPS, and nor was the amplitude of “tool” preferences in the face-preferring area of the fusiform gyrus.

The findings of Garcea et al. (2019) motivate what we termed “domain-specific diaschisis” (a form of “dynamic diaschisis,” see Price et al., 2001). The VLAM analyses reported by Garcea and colleagues indicate that tool responses in the ventral stream are *dependent* on the integrity of processing in the left aIPS, which is exactly what would be predicted if tool responses in that ventral stream region are *driven* by the intersection of two constraints—one coming through the ventral visual hierarchy and one coming from aIPS via the dorsal stream.

A complementary study by Lee et al. (2019) showed that TDCS to the parietal cortex transiently modulated neural responses to (specifically) tools in (specifically) the medial fusiform gyrus. An important study that should be carried out would be to pair TMS *during* fMRI to test whether concurrent stimulation to parietal tool-preferring regions up- or downregulates responses (specifically) to tools in (specifically) the medial fusiform gyrus and collateral sulcus, as well as the left posterior middle temporal gyrus.

Reframing expectations for a domain-specific account of occipito-temporal cortex organization

The distributed domain-specific hypothesis combines the “old idea” that the brain comes preprogrammed with a limited number of learning and information processing systems oriented toward fundamentally different computational problems (Spelke and Kinzler, 2007) with an account of the causes of neural specificity for different semantic categories in occipito-temporal (and other) brain areas. Examples of computational domains could be: inferring and thinking about the inner mental lives of other agents, navigation and context-dependent memory, and physical reasoning about how first person actions will change the state of the world (Mahon, 2020). Different computational goals define “successful” processing of items from different categories: Successfully processing geographical landmarks in the context of navigation is a fundamentally different process from inferring someone’s motivation for an action, or how an object can be manipulated to accomplish a behavioral goal (see discussion of computational “eccentricity” by Fodor, 1983).

It is convenient, if somewhat misleading, shorthand to say that a network “is domain-specific for tools” (or faces, animals, places, words, body parts, numbers). “Tools” are not the “domain”—they are a way of defining a set of “the things in the world” that form an equivalence class with respect to a given hypothesized learning and information processing system. It follows that the goal of defining what the relevant semantic categories are is not about picking out a natural kind in the world (or in the mind/brain). A “category” as such has no objective validity. Defining “categories” operationally to measure “category-specificity” in the brain is about individuating the right set of computations to be able to experimentally interrogate them. The burden of a domain-specific theory thus shifts from showing “category selectivity” in the response profiles of individual regions to articulating (i) the different computational problems that define separable domain-specific systems, (ii) their constituent regions, and (iii) and

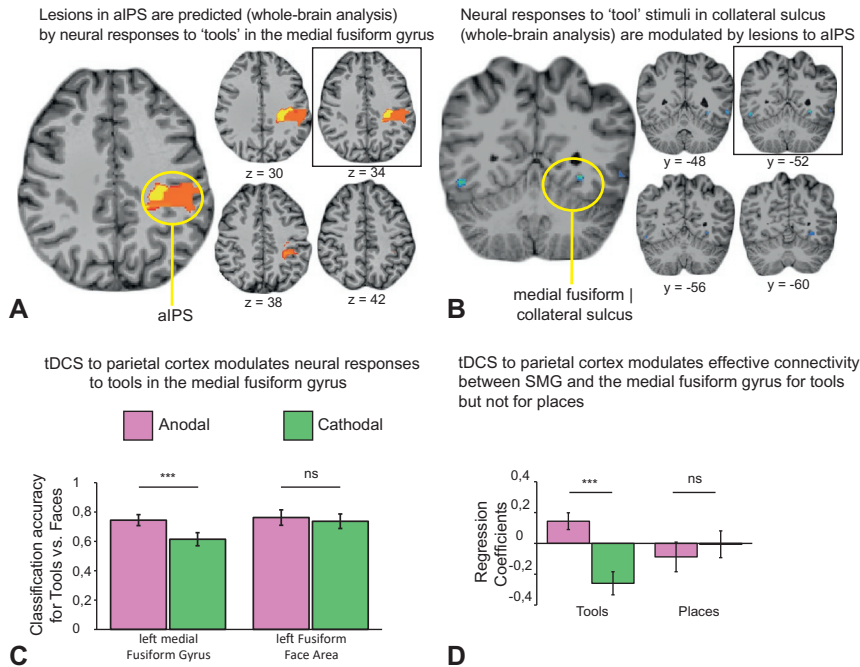


Fig. 13.4. Causal evidence that neural responses to manipulable objects in occipito-temporal cortex depend on real-time processing in the parietal cortex. Two recent findings provide causal evidence for the hypothesis that processing of manipulable objects in the ventral stream is modulated by parietal action representations. Garcea et al. (2019) found that lesions involving aIPS were associated with *reduced* fMRI responses to tool stimuli in medial ventral temporal areas; however, lesions to aIPS were not associated with reduced responses to place stimuli in the same regions (A). This is despite the fact that, if anything, neural responses in medial ventral occipito-temporal areas were stronger for place than tool stimuli. A whole-brain analysis that searched for where BOLD activity was inversely related to the probability of a lesion to aIPS, identified the medial fusiform gyrus and collateral sulcus (B; this analysis also identified the posterior inferior/middle temporal gyrus). In a separate study, in healthy participants, Lee et al. (2019) found that cathodal tDCS to the left parietal cortex disrupted voxel-wise pattern discriminability between tools and animals (C) in medial ventral occipito-temporal cortex, but not in lateral aspects of the fusiform gyrus (face-preferring area). Neural similarity among tool stimuli was increased after excitatory (anodal) stimulation of the left parietal cortex and reduced after inhibitory (cathodal) stimulation of the left parietal cortex. Moreover, as shown in (D), cathodal vs anodal stimulation modulated effective functional connectivity between SMG and the medial fusiform gyrus for tool, but not for place, stimuli. These findings indicate that parietal representations of object-directed action causally modulate, online, visual object processing of graspable objects in the occipital-temporal cortex.

how they connect, and ultimately (iv) the dynamics that govern their interactions as processing enables behavioral goals.

If a region exhibits “category-specificity,” that means it responds maximally or selectively to one category of items compared to others. Domain-specificity is not about the things in the world that engage a process; domain-specificity is about the problem that is being solved by a whole network of regions working together. The scope of any particular process, implemented by any given region, dictates the types of stimuli that will be successfully operated upon by that process. The processes supported by separable regions of a (hypothesized) domain-specific system fit together in the service of the computational goal that defines that domain.

This leads to the conclusion: Domain-specificity is not the same thing as category-specificity. Some regions

exhibiting category-specificity are part of domain-specific networks. But regions that are not “category-specific” may nonetheless participate in supporting domain-specific computational goals. It follows that the test of the domain-specificity of a region is not reducible to a test of the selectivity of a region to one or another category. That approach is a legacy of defining domain-specificity *in terms of* category-selectivity. The response profile of a region will be as sharp as the boundary that defines the process implemented by that region. This is both a claim about what domain-specificity *is*, and a methodological guardrail about how we may go about testing hypotheses of domain-specificity.

Consider the following analogy. The digestive system is “domain-specific” for digestion. The stomach is specialized for a set of processes, and because of what types of inputs those processes operate over, one can say that

the “processes of the stomach are specialized for food.” And yet, if one swallows a nonfood item (say, chalk or some pebbles), the stomach will “respond”—it will secrete digestive juices and “behave as if” what was swallowed was food. What this means is just that the stomach (on its own) does not have a “food detector” that gates its responses according to whether what was ingested was food. The analogy to digestion isolates the intuition that whether or not the stomach has a “food detector” is a completely different issue from whether the stomach is specialized for digesting food. The hypothesis “*the stomach is specialized for food*” is not embarrassed by the fact that the stomach “responds” to chalk. Support for the hypothesis “*the stomach is specialized for food*” would consist of showing that the stomach has a well-defined set of operations that it runs when it gets an input; and, when those operations are run over inputs that are food, those operations result in outputs from the stomach that are well-defined inputs to the next organ in the processing chain. When the inputs to the stomach are not food, then those operations do not generate useful outputs (even though they may “try” to run their processes). Thus, the test of a theory of the process specificity of the stomach is not evaluable solely based on how the stomach *responds*, or whether it responds in a graded manner to nonfood items. The test of such a theory amounts to whether the rest of the “digestive network” can handle the outputs of the stomach, according to the inputs to the stomach (and potentially the state of the stomach when it got those inputs).

The brain regions in the schematic of [Fig. 13.1](#) exhibit preferences for manipulable objects. “Preference” is an intentionally vague term intended to subsume different signatures of process specificity: amplitude, univariate contrasts, repetition suppression and adaptation, structural connectivity, alignment of connectivity with stimulus preferences, multivariate pattern analysis, and so on. Consider the pattern of univariate contrast effects observed across three regions highlighted in [Fig. 13.1](#) ([Culham et al., 2003](#); [Mahon et al., 2007](#)): The supramarginal gyrus, the dorsal occipital cortex, and the medial fusiform gyrus. The left inferior parietal lobule responds more to manipulable objects (hammer) compared to *both* large nonmanipulable objects (refrigerator) as well as small graspable objects that do not have a stereotyped manner of manipulation (wallet). By contrast, caudal IPS and dorsal occipital cortex (V6A; [Pitzalis et al., 2013](#)) not only supports the computation of reach trajectories based on volumetric analysis of targets, but also responds to any graspable target (hammer, fork, wallet, banana; [Fang and He, 2005](#)). Finally, the medial fusiform gyrus, which has been demonstrated to enjoy privileged functional connectivity to inferior parietal areas involved

in grasping, and which processes surface texture information, responds more to manipulable objects than to faces and animals. And yet, that same medial fusiform region may respond more strongly to large nonmanipulable objects and place stimuli ([Downing et al., 2006](#); [Mahon et al., 2007](#)) than to manipulable objects.

The reason that one region (in a complex network) exhibits preferences for manipulable objects may be very different from the reason why another region in the same domain-specific network exhibits such preferences. Why? The two regions implement different processes—an uncontroversial premise. This is a feature of domain-specificity, not a bug that needs to be addressed. A hammer may drive responses in both the supramarginal gyrus and dorsal occipital cortex—not because it is a “tool,” but because it is manipulable (*vis-à-vis* the supramarginal gyrus) and graspable (*vis-à-vis* the dorsal occipital cortex). Showing that dorsal occipital cortex responds to a banana while the supramarginal gyrus does not provides important clues as to the separable computations implemented by those different regions—and importantly serves to triangulate the types of broad computational goals that may be attributed to the network in question.

This way of framing domain-specificity departs from the way it has been discussed over the past two decades in the context of functional MRI studies of category-specificity. The field has been driven by the following dynamic: A region is found that is specialized for processing category X (for instance, that an area in the fusiform gyrus is specialized for processing faces). Other researchers show that the region in question also responds to non-X’s (e.g., greebels), and argue (correctly) that the theory that stated that region A is specialized for Xs has been falsified ([Gauthier et al., 1999](#)). But if we step back, the inferential landscape is very different: The conclusion that region A is not specialized for X’s because it responds to non-X’s is no more solid than the conclusion that the stomach is not specialized for food because it will “respond” to nonfood substances (e.g., chalk). The “mistake” was not in the critique of domain-specificity; it was in the initial claim that region A is specialized for category X. It would have been more accurate to state the hypothesis as: region A is specialized for computation Y, and when computation Y is applied to X’s, it generates outputs that are useful to other systems/computations, but this not the case when the computation is applied to non-X’s.

CONCLUSIONS

There are many nontrivial parallels between phenomena of category-specificity in the neuropsychological and functional neuroimaging literatures. First, there are broad

divisions of labor within visual processing that are optimized for different applications of visual information to behavioral goals (concurrent actions, recognition, perceptual constancy for identification). Second, the categories that emerge from the neuropsychological literature largely map onto the categories that emerge in functional neuroimaging: faces, animals, tools, printed words, places, body parts, numbers. Third, the existence of category-specificity in neural organization as revealed by functional neuroimaging and neuropsychology does not seem reducible to sensory-modality-based principles of organization. Many proposals have been made about dimensions that correlate with semantic category distinctions; work over the past several decades argues for the conclusion, at least in this telling of the history, that those proposals each fall short in accounting for the full pattern of dissociations observed in behavior and in the brain. Fourth, there is evidence that the phenomena of category-specificity are present independent of substantial variability in sensory experience. Fifth, evidence from twin studies indicates that core aspects of functional and neural organization are genetically constrained. Sixth, evidence from selective rearing studies indicates a critical period of inputs is required for some patterns of neural organization to emerge. Seventh, functional and structural connectivity is a key factor for understanding the anatomical organization of category-preferences, and by hypothesis, the macroscopic organization of knowledge in the brain.

The foregoing considerations motivate an approach that emphasizes connectivity-based constraints as providing the initial, and primary, scaffolding that drives the organization of visual recognition and conceptual knowledge about the world (Mahon et al., 2009; Mahon and Caramazza, 2011). If one takes connectivity as the *starting point* for understanding regional specialization in the ventral stream, then the traditional way of thinking about the causes of category-specificity in the ventral stream are flipped around. Priority is placed on understanding how ventral stream regions interact with regions outside of the ventral stream. Ultimately, the value of this broader proposal will be weighed in its ability to generate new predictions and, at a pragmatic level, whether it serves as a useful paradigm with which to study how the brain works.

ACKNOWLEDGMENTS

The ideas about domain-specificity described herein have been developed collaboratively with Alfonso Caramazza over the past two decades of discussion. I am grateful to Gabriele Miceli and Paolo Bartolomeo

for constructive feedback on an earlier draft of this chapter, and to Kathleen Haaland and David Plaut for feedback on earlier formulations of the review. The preparation of this manuscript was supported in part by Grants R01NS089069 and R01EY028535.

REFERENCES

- Abbasi N, Duncan J, Rajimehr R (2020). Genetic influence is linked to cortical morphology in category-selective areas of visual cortex. *Nat Commun* 11: 709. <https://doi.org/10.1038/s41467-020-14610-8>.
- Allison T, McCarthy G, Nobre A et al. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cereb Cortex* 4: 544–554.
- Amedi A, Malach R, Hendler T et al. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nat Neurosci* 4: 324–330. <https://doi.org/10.1038/85201>.
- Amedi A, Stern W, Camprodon J et al. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nat Neurosci* 10: 687–689. <https://doi.org/10.1038/nn1912>.
- Arcaro M, Livingstone M (2017). A hierarchical, retinotopic proto-organization of the primate visual system at birth. *eLife* 6. <https://doi.org/10.7554/eLife.26196>.
- Arcaro MJ, Schade PF, Vincent JL et al. (2017). Seeing faces is necessary for face-domain formation. *Nat Neurosci* 20: 1404–1412. <https://doi.org/10.1038/nn.4635>.
- Avidan G, Harel M, Hendler T et al. (2002). Contrast sensitivity in human visual areas and its relationship to object recognition. *J Neurophysiol* 87: 3102–3116. <https://doi.org/10.1152/Jn.00669.2001>.
- Bar M, Kassam K, Ghuman A et al. (2006). Top-down facilitation of visual recognition. *Proc Natl Acad Sci U S A* 103: 449–454. <https://doi.org/10.1073/pnas.0507062103>.
- Barr M, Aminoff E (2003). Cortical analysis of visual context. *Neuron* 38: 347–358.
- Bates E, Wilson S, Saygin A et al. (2003). Voxel-based lesion-symptom mapping. *Nat Neurosci* 6: 448–450. <https://doi.org/10.1038/nn1050>.
- Beauchamp MS, Lee KE, Haxby JV et al. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron* 24: 149–159.
- Beauchamp MS, Lee KE, Haxby JV et al. (2003). fMRI responses to video and point-light displays of moving humans and manipulable objects. *J Cogn Neurosci* 15: 991–1001.
- Binkofski F, Dohle C, Posse S et al. (1998). Human anterior intraparietal area subserves prehension: a combined lesion and functional MRI activation study. *Neurology* 50: 1253–1259. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/9595971>.
- Blundo C, Ricci M, Miller L (2006). Category-specific knowledge deficit for animals in a patient with herpes simplex encephalitis. *Cogn Neuropsychol* 23: 1248–1268.
- Bouhali F, Thiebaut de Schotten M, Pinel P et al. (2014). Anatomical connections of the visual word form area.

- J Neurosci 34: 15402–15414. <https://doi.org/10.1523/JNEUROSCI.4918-13.2014>.
- Bracci S, Peelen M (2013). Body and object effectors: the organization of object representations in high-level visual cortex reflects body-object interactions. *J Neurosci* 33: 18247–18258. <https://doi.org/10.1523/JNEUROSCI.1322-13.2013>.
- Bracci S, Ietswaart M, Peelen MV et al. (2010). Dissociable neural responses to hands and non-hand body parts in human left extrastriate visual cortex. *J Neurophysiol* 103: 3389–3397. <https://doi.org/10.1152/jn.00215.2010>.
- Bracci S, Cavina-Pratesi C, Ietswaart M et al. (2012). Closely overlapping responses to tools and hands in left lateral occipitotemporal cortex. *J Neurophysiol* 107: 1443–1456. <https://doi.org/10.1152/jn.00619.2011>.
- Buchel C, Price C, Friston K (1998). A multimodal language region in the ventral visual pathway. *Nature* 394: 274–277. <https://doi.org/10.1038/28389>.
- Cant JS, Goodale MA (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb Cortex* 17: 713–731. <https://doi.org/10.1093/cercor/bhk022>.
- Capitani E, Laiacona M, Mahon B et al. (2003). What are the facts of semantic category-specific deficits? A critical review of the clinical evidence. *Cogn Neuropsychol* 20: 213–261. <https://doi.org/10.1080/02643290244000266>.
- Caramazza A, Mahon BZ (2003). The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends Cogn Sci* 7: 354–361. [https://doi.org/10.1016/s1364-6613\(03\)00159-1](https://doi.org/10.1016/s1364-6613(03)00159-1).
- Caramazza A, Shelton JR (1998). Domain specific knowledge systems in the brain: the animate-inanimate distinction. *J Cogn Neurosci* 10: 1–34.
- Carroll J, Conway BR (2021). Color vision. *Handb Clin Neurol* 178: 131–153. <https://doi.org/10.1016/B978-0-12-821377-3.00005-2>.
- Cavina-Pratesi C, Kentridge RW, Heywood CA et al. (2010). Separate processing of texture and form in the ventral stream: evidence from fMRI and visual agnosia. *Cereb Cortex* 20: 433–446. <https://doi.org/10.1093/cercor/bhp111>.
- Changizi MA, Zhang Q, Shimojo S (2006). Bare skin, blood and the evolution of primate colour vision. *Biol Lett* 2: 217–221. <https://doi.org/10.1098/rsbl.2006.0440>.
- Chao LL, Martin A (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage* 12: 478–484. <https://doi.org/10.1006/nimg.2000.0635>.
- Chao L, Haxby J, Martin A (1999a). Attribute-based neural substrates in posterior temporal cortex for perceiving and knowing about objects. *Nat Neurosci* 2: 913–919.
- Chao LL, Haxby JV, Martin A (1999b). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci* 2: 913–919. <https://doi.org/10.1038/13217>.
- Chen L, Rogers TT (2015). A model of emergent category-specific activation in the posterior fusiform gyrus of sighted and congenitally blind populations. *J Cogn Neurosci* 27: 1981–1999. https://doi.org/10.1162/jocn_a_00834.
- Chen Q, Garcea F, Almeida J et al. (2017). Connectivity-based constraints on category-specificity in the ventral object processing pathway. *Neuropsychologia* 105: 184–196. <https://doi.org/10.1016/j.neuropsychologia.2016.11.014>.
- Chen J, Snow JC, Culham JC et al. (2018). What role does “elongation” play in “tool-specific” activation and connectivity in the dorsal and ventral visual streams? *Cereb Cortex* 28: 1117–1131. <https://doi.org/10.1093/cercor/bhx017>.
- Cohen L, Dehaene S, Naccache L et al. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123: 291–307.
- Conway BR (2018). The organization and operation of inferior temporal cortex. *Annu Rev Vis Sci* 4: 381–402. <https://doi.org/10.1146/annurev-vision-091517-034202>.
- Culham JC, Danckert SL, DeSouza JF et al. (2003). Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Exp Brain Res* 153: 180–189. <https://doi.org/10.1007/s00221-003-1591-5>.
- Dehaene S, Cohen L (2007). Cultural recycling of cortical maps. *Neuron* 56: 384–398. <https://doi.org/10.1016/j.neuron.2007.10.004>.
- Dehaene S, Cohen L, Sigman M et al. (2005). The neural code for written words: a proposal. *Trends Cogn Sci* 9: 335–341. <https://doi.org/10.1016/j.tics.2005.05.004>.
- Downing PE, Jiang YH, Shuman M et al. (2001). A cortical area selective for visual processing of the human body. *Science* 293: 2470–2473. <https://doi.org/10.1126/science.1063414>.
- Downing P, Chan A, Peelen M et al. (2006). Domain specificity in visual cortex. *Cereb Cortex* 16: 1453–1461. <https://doi.org/10.1093/cercor/bhj086>.
- Duchaine BC, Nakayama K (2006). Developmental prosopagnosia: a window to content-specific face processing. *Curr Opin Neurobiol* 16: 166–173. <https://doi.org/10.1016/j.conb.2006.03.003>.
- Dundas EM, Plaut DC, Behrmann M (2013). The joint development of hemispheric lateralization for words and faces. *J Exp Psychol Gen* 142: 348–358. <https://doi.org/10.1037/a0029503>.
- Epstein R, Kanwisher N (1998). A cortical representation of the local visual environment. *Nature* 392: 598–601.
- Fang F, He S (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nat Neurosci* 8: 1380–1385. <https://doi.org/10.1038/nn1537>.
- Farah MJ, McClelland JL (1991). A computational model of semantic memory impairment: modality specificity and emergent category specificity. *J Exp Psychol Gen* 120: 339–357. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/1837294>.
- Farah M, Rabinowitz C (2003). Genetic and environmental influences on the organization of semantic memory in the brain: is “living things” an innate category? *Cogn Neuropsychol* 20: 401–408.

- Fodor J (1983). *The modularity of mind: an essay in faculty psychology*, The MIT Press, Cambridge MA.
- Friedman-Hill SR, Robertson LC, Treisman A (1995). Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions. *Science* 269: 853–855. <https://doi.org/10.1126/science.7638604>.
- Friston KJ, Rotshtein P, Geng JJ et al. (2006). A critique of functional localisers. *Neuroimage* 30: 1077–1087. <https://doi.org/10.1016/j.neuroimage.2005.08.012>.
- Gallivan JP, Chapman CS, McLean DA et al. (2013a). Activity patterns in the category-selective occipitotemporal cortex predict upcoming motor actions. *Eur J Neurosci* 38: 2408–2424. <https://doi.org/10.1111/ejn.12215>.
- Gallivan JP, McLean DA, Valyear KF et al. (2013b). Decoding the neural mechanisms of human tool use. *eLife* 2: e00425. <https://doi.org/10.7554/eLife.00425>.
- Gallivan JP, Cant JS, Goodale MA et al. (2014). Representation of object weight in human ventral visual cortex. *Curr Biol* 24: 1866–1873. <https://doi.org/10.1016/j.cub.2014.06.046>.
- Garcea F, Mahon B (2014). Parcellation of left parietal tool-representations by functional connectivity. *Neuropsychologia* 60: 131–143. <https://doi.org/10.1016/j.neuropsychologia.2014.05.018>.
- Garcea F, Chen Q, Vargas R et al. (2018). Task- and domain-specific modulation of functional connectivity in the ventral and dorsal object-processing pathways. *Brain Struct Funct*. <https://doi.org/10.1007/s00429-018-1641-1>.
- Garcea F, Almeida J, Sims M 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. <https://doi.org/10.1093/cercor/bhy183>.
- Gauthier I, Tarr MJ, Anderson AW et al. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat Neurosci* 2: 568–573. <https://doi.org/10.1038/9224>.
- Geskin J, Behrmann M (2018). Congenital prosopagnosia without object agnosia? A literature review. *Cogn Neuropsychol* 35: 4–54. <https://doi.org/10.1080/02643294.2017.1392295>.
- Goldenberg G, Hagmann S (1998). Tool use and mechanical problem solving in apraxia. *Neuropsychologia* 36: 581–589. [https://doi.org/10.1016/s0028-3932\(97\)00165-6](https://doi.org/10.1016/s0028-3932(97)00165-6).
- Gould S, Lewontin R (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc R Soc Lond B Biol Sci* 205: 581–598.
- Grill-Spector K, Malach R (2004). The human visual cortex. *Annu Rev Neurosci* 27: 649–677. <https://doi.org/10.1146/annurev.neuro.27.070203.144220>.
- Hasson U, Levy I, Behrmann M et al. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34: 479–490. [https://doi.org/10.1016/s0896-6273\(02\)00662-1](https://doi.org/10.1016/s0896-6273(02)00662-1).
- Hasson U, Avidan G, Deouell L et al. (2003). Face-selective activation in a congenital prosopagnosic subject. *J Cogn Neurosci* 15: 419–431.
- Haxby JV, Gobbini MI, Furey ML et al. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293: 2425–2430. <https://doi.org/10.1126/science.1063736>.
- He C, Peelen MV, Han Z et al. (2013). Selectivity for large nonmanipulable objects in scene-selective visual cortex does not require visual experience. *Neuroimage* 79: 1–9. <https://doi.org/10.1016/j.neuroimage.2013.04.051>.
- Heider F, Simmel M (1944). An experimental study of apparent behavior. *Am J Psychol* 57: 243–259.
- Hutchison RM, Gallivan JP, Culham JC et al. (2012). Functional connectivity of the frontal eye fields in humans and macaque monkeys investigated with resting-state fMRI. *J Neurophysiol* 107: 2463–2474. <https://doi.org/10.1152/jn.00891.2011>.
- James TW, Humphrey GK, Gati JS et al. (2002). Differential effects of viewpoint on object-driven activation in dorsal and ventral streams. *Neuron* 35: 793–801.
- Johnson MH, Dziurawiec S, Ellis H et al. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition* 40: 1–19. [https://doi.org/10.1016/0010-0277\(91\)90045-6](https://doi.org/10.1016/0010-0277(91)90045-6).
- Kamps FS, Hendrix CL, Brennan PA et al. (2020). Connectivity at the origins of domain specificity in the cortical face and place networks. *Proc Natl Acad Sci U S A* 117: 6163–6169. <https://doi.org/10.1073/pnas.1911359117>.
- Kanwisher N, McDermott J, Chun MM (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17: 4302–4311. Retrieved from <http://www.jneurosci.org/content/jneuro/17/11/4302.full.pdf>.
- Kemmerer D, Gonzalez-Castillo J (2010). The Two-Level Theory of verb meaning: an approach to integrating the semantics of action with the mirror neuron system. *Brain Lang* 112: 54–76. <https://doi.org/10.1016/j.bandl.2008.09.010>.
- Konkle T, Oliva A (2012). A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74: 1114–1124. <https://doi.org/10.1016/j.neuron.2012.04.036>.
- Kriegeskorte N, Mur M, Ruff DA et al. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60: 1126–1141. <https://doi.org/10.1016/j.neuron.2008.10.043>.
- Lafer-Sousa R, Conway BR (2013). Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nat Neurosci* 16: 1870–1878. <https://doi.org/10.1038/nn.3555>.
- Lafer-Sousa R, Conway BR, Kanwisher NG (2016). Color-biased regions of the ventral visual pathway lie between face- and place-selective regions in humans, as in macaques. *J Neurosci* 36: 1682–1697. <https://doi.org/10.1523/JNEUROSCI.3164-15.2016>.
- Lee D, Mahon BZ, Almeida J (2019). Action at a distance on object-related ventral temporal representations. *Cortex* 117: 157–167. <https://doi.org/10.1016/j.cortex.2019.02.018>.
- Levy I, Hasson U, Avidan G et al. (2001). Center-periphery organization of human object areas. *Nat Neurosci* 4: 533–539. <https://doi.org/10.1038/87490>.

- Li J, Osher DE, Hansen HA et al. (2020). Innate connectivity patterns drive the development of the visual word form area. *Sci Rep* 10: 18039. <https://doi.org/10.1038/s41598-020-75015-7>.
- Livingstone M, Hubel D (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* 240: 740–749. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/3283936>.
- Mahon B (2020). The representation of tools in the human brain. In: D Poeppel, M Gazzaniga (Eds.), *The new cognitive neurosciences*, sixth edn. MIT Press, Cambridge, MA.
- Mahon BZ, Caramazza A (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annu Rev Psychol* 60: 27–51. <https://doi.org/10.1146/annurev.psych.60.110707.163532>.
- Mahon B, Caramazza A (2011). What drives the organization of object knowledge in the brain? *Trends Cogn Sci* 15: 97–103. <https://doi.org/10.1016/j.tics.2011.01.004>.
- Mahon B, Milleville S, Negri G et al. (2007). Action-related properties shape object representations in the ventral stream. *Neuron* 55: 507–520. <https://doi.org/10.1016/j.neuron.2007.07.011>.
- Mahon BZ, Anzellotti S, Schwarzbach J et al. (2009). Category-specific organization in the human brain does not require visual experience. *Neuron* 63: 397–405. <https://doi.org/10.1016/j.neuron.2009.07.012>.
- Mahon BZ, Schwarzbach J, Caramazza A (2010). The representation of tools in left parietal cortex is independent of visual experience. *Psychol Sci* 21: 764–771. <https://doi.org/10.1177/0956797610370754>.
- Mahon BZ, Kumar N, Almeida J (2013). Spatial frequency tuning reveals interactions between the dorsal and ventral visual systems. *J Cogn Neurosci* 25: 862–871. https://doi.org/10.1162/jocn_a_00370.
- Martin A (2006). Shades of Dejerine—forging a causal link between the visual word form area and reading. *Neuron* 50: 173–175. <https://doi.org/10.1016/J.Neuron.2006.04.004>.
- Martin A (2007). The representation of object concepts in the brain. *Annu Rev Psychol* 58: 25–45.
- 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. <https://doi.org/10.3758/s13423-015-0842-3>.
- Martin A, Weisberg J (2003). Neural foundations for understanding social and mechanical concepts. *Cogn Neuropsychol* 20: 575–587.
- Mechelli A, Sartori G, Orlandi P et al. (2006). Semantic relevance explains category effects in medial fusiform gyri. *Neuroimage* 3: 992–1002.
- Miceli G, Capasso R, Daniele A et al. (2000). Selective deficit for people's names following left temporal damage: an impairment of domain-specific conceptual knowledge. *Cogn Neuropsychol* 17: 489–516.
- Miceli G, Fouch E, Capasso R et al. (2001). The dissociation of color from form and function knowledge. *Nat Neurosci* 4: 662–667. <https://doi.org/10.1038/88497>.
- Mruczek RE, von Loga IS, Kastner S (2013). The representation of tool and non-tool object information in the human intraparietal sulcus. *J Neurophysiol* 109: 2883–2896. <https://doi.org/10.1152/jn.00658.2012>.
- Munoz-Rubke F, Olson D, Will R et al. (2018). Functional fixedness in tool use: learning modality, limitations and individual differences. *Acta Psychol (Amst)* 190: 11–26.
- Nasr S, Echavarria CE, Tootell RB (2014). Thinking outside the box: rectilinear shapes selectively activate scene-selective cortex. *J Neurosci* 34: 6721–6735. <https://doi.org/10.1523/JNEUROSCI.4802-13.2014>.
- Natu VS, Arcaro MJ, Barnett MA et al. (2021). Sulcal depth in the medial ventral temporal cortex predicts the location of a place-selective region in macaques, children, and adults. *Cereb Cortex* 31: 48–61. <https://doi.org/10.1093/cercor/bhaa203>.
- Op de Beeck HP, Haushofer J, Kanwisher NG (2008). Interpreting fMRI data: maps, modules and dimensions. *Nat Rev Neurosci* 9: 123–135. <https://doi.org/10.1038/nrn2314>.
- Osher DE, Saxe RR, Koldewyn K et al. (2016). Structural connectivity fingerprints predict cortical selectivity for multiple visual categories across cortex. *Cereb Cortex* 26: 1668–1683. <https://doi.org/10.1093/cercor/bhu303>.
- Peelen MV, Bracci S, Lu X et al. (2013). Tool selectivity in left occipitotemporal cortex develops without vision. *J Cogn Neurosci* 25: 1225–1234. https://doi.org/10.1162/jocn_a_00411.
- Pitcher D, Charles L, Devlin JT et al. (2009). Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Curr Biol* 19: 319–324. <https://doi.org/10.1016/j.cub.2009.01.007>.
- Pitzalis S, Sereno MI, Committeri G et al. (2013). The human homologue of macaque area V6A. *Neuroimage* 82: 517–530. <https://doi.org/10.1016/j.neuroimage.2013.06.026>.
- Plaut DC, Behrmann M (2011). Complementary neural representations for faces and words: a computational exploration. *Cogn Neuropsychol* 28: 251–275. <https://doi.org/10.1080/02643294.2011.609812>.
- Poeppel D (2012). The maps problem and the mapping problem: two challenges for a cognitive neuroscience of speech and language. *Cogn Neuropsychol* 29: 34–55. <https://doi.org/10.1080/02643294.2012.710600>.
- Polk TA, Park J, Smith MR et al. (2007). Nature versus nurture in ventral visual cortex: a functional magnetic resonance imaging study of twins. *J Neurosci* 27: 13921–13925. <https://doi.org/10.1523/JNEUROSCI.4001-07.2007>.
- Prentiss EK, Schneider CL, Williams ZR et al. (2018). Spontaneous in-flight accommodation of hand orientation to unseen grasp targets: a case of action blindsight. *Cogn Neuropsychol* 35: 343–351. <https://doi.org/10.1080/02643294.2018.1432584>.
- Price C, Warburton E, Moore C et al. (2001). Dynamic diaschisis: anatomically remote and context-sensitive human brain lesions. *J Cogn Neurosci* 13: 419–429. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11388916.

- Riesenhuber M (2007). Appearance isn't everything: news on object representation in cortex. *Neuron* 55: 341–344. <https://doi.org/10.1016/j.neuron.2007.07.017>.
- Riesenhuber M, Poggio T (1999). Hierarchical models of object recognition in cortex. *Nat Neurosci* 2: 1019–1025. <https://doi.org/10.1038/14819>.
- Rogers TT, Hocking J, Mechelli A et al. (2005). Fusiform activation to animals is driven by the process, not the stimulus. *J Cogn Neurosci* 17: 434–445.
- Rothi LJG, Ochipa C, Heilman KM (1991). A cognitive neuropsychological model of limb praxis. *Cogn Neuropsychol* 8: 443–458. <https://doi.org/10.1080/02643299108253382>.
- Saxe R, Brett M, Kanwisher N (2006). Divide and conquer: a defense of functional localizers. *Neuroimage* 30: 1088–1096. discussion 1097–1099. <https://doi.org/10.1016/j.neuroimage.2005.12.062>.
- Saygin ZM, Osher DE, Koldewyn K et al. (2011). Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. *Nat Neurosci* 15: 321–327. <https://doi.org/10.1038/nn.3001>.
- Saygin ZM, Osher DE, Norton ES et al. (2016). Connectivity precedes function in the development of the visual word form area. *Nat Neurosci* 19: 1250–1255. <https://doi.org/10.1038/nn.4354>.
- Simmons WK, Ramjee V, Beauchamp MS et al. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia* 45: 2802–2810. <https://doi.org/10.1016/j.neuropsychologia.2007.05.002>.
- Siuda-Krzywicka K, Bartolomeo P (2020). What cognitive neurology teaches us about our experience of color. *Neuroscientist* 26: 252–265. <https://doi.org/10.1177/1073858419882621>.
- Siuda-Krzywicka K, Witzel C, Chabani E et al. (2019). Color categorization independent of color naming. *Cell Rep* 28. <https://doi.org/10.1016/j.celrep.2019.08.003>. 2471–2479.e5.
- Siuda-Krzywicka K, Witzel C, Taga M et al. (2020). When colours split from objects: the disconnection of colour perception from colour language and colour knowledge. *Cogn Neuropsychol* 37: 325–339. <https://doi.org/10.1080/02643294.2019.1642861>.
- Siuda-Krzywicka K, Witzel C, Bartolomeo P et al. (2021). Color naming and categorization depend on distinct functional brain networks. *Cereb Cortex* 31: 1106–1115. <https://doi.org/10.1093/cercor/bhaa278>.
- Spelke ES, Kinzler KD (2007). Core knowledge. *Dev Sci* 10: 89–96. <https://doi.org/10.1111/j.1467-7687.2007.00569.x>.
- Srihasam K, Vincent JL, Livingstone MS (2014). Novel domain formation reveals proto-architecture in inferotemporal cortex. *Nat Neurosci* 17: 1776–1783. <https://doi.org/10.1038/nn.3855>.
- Stassenko A, Garcea FE, Dombovy M et al. (2014). When concepts lose their color: a case of object-color knowledge impairment. *Cortex* 58: 217–238. <https://doi.org/10.1016/j.cortex.2014.05.013>.
- Stevens WD, Tessler MH, Peng CS et al. (2015). Functional connectivity constrains the category-related organization of human ventral occipitotemporal cortex. *Hum Brain Mapp* 36: 2187–2206. <https://doi.org/10.1002/hbm.22764>.
- Stoerig P, Cowey A (1989). Wavelength sensitivity in blindsight. *Nature* 342: 916–918. <https://doi.org/10.1038/342916a0>.
- Striem-Amit E, Cohen L, Dehaene S et al. (2012). Reading with sounds: sensory substitution selectively activates the visual word form area in the blind. *Neuron* 76: 640–652. <https://doi.org/10.1016/j.neuron.2012.08.026>.
- Thomas C, Avidan G, Humphreys K et al. (2009). Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nat Neurosci* 12: 29–31. <https://doi.org/10.1038/nn.2224>.
- Tsao DY, Freiwald WA, Tootell RB et al. (2006). A cortical region consisting entirely of face-selective cells. *Science* 311: 670–674. <https://doi.org/10.1126/science.1119983>.
- Valyear KF, Culham JC (2010). Observing learned object-specific functional grasps preferentially activates the ventral stream. *J Cogn Neurosci* 22: 970–984. <https://doi.org/10.1162/jocn.2009.21256>.
- Valyear KF, Gallivan JP, McLean DA et al. (2012). fMRI repetition suppression for familiar but not arbitrary actions with tools. *J Neurosci* 32: 4247–4259. <https://doi.org/10.1523/JNEUROSCI.5270-11.2012>.
- Wilmer JB, Germine L, Chabris CF et al. (2010). Human face recognition ability is specific and highly heritable. *Proc Natl Acad Sci U S A* 107: 5238–5241. <https://doi.org/10.1073/pnas.0913053107>.
- Wu W (2008). Visual attention, conceptual content, and doing it right. *Mind* 117: 1003–1033.
- Wurm MF, Caramazza A, Lingnau A (2017). Action categories in lateral occipitotemporal cortex are organized along sociality and transitivity. *J Neurosci* 37: 562–575. <https://doi.org/10.1523/JNEUROSCI.1717-16.2016>.
- Zeki S, Ffytche DH (1998). The Riddoch syndrome: insights into the neurobiology of conscious vision. *Brain* 121: 25–45. <https://doi.org/10.1093/brain/121.1.25>.
- Zhang Y, Kimberg DY, Coslett HB et al. (2014). Multivariate lesion-symptom mapping using support vector regression. *Hum Brain Mapp* 35: 5861–5876. <https://doi.org/10.1002/hbm.22590>.
- Zhu Q, Song Y, Hu S et al. (2010). Heritability of the specific cognitive ability of face perception. *Curr Biol* 20: 137–142. <https://doi.org/10.1016/j.cub.2009.11.067>.
- Zhuang C, Yan S, Nayebi A et al. (2021). Unsupervised neural network models of the ventral visual stream. *Proc Natl Acad Sci U S A* 118. <https://doi.org/10.1073/pnas.2014196118>.