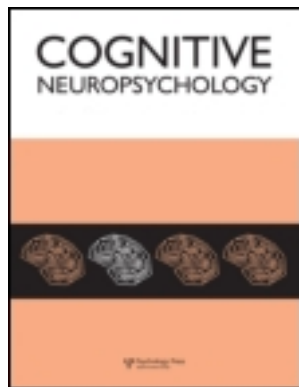


This article was downloaded by: [Villanova University]

On: 25 January 2012, At: 06:03

Publisher: Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Cognitive Neuropsychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pcgn20>

### Complementary neural representations for faces and words: A computational exploration

David C. Plaut<sup>a</sup> & Marlene Behrmann<sup>a</sup>

<sup>a</sup> Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA

Available online: 20 Dec 2011

To cite this article: David C. Plaut & Marlene Behrmann (2011): Complementary neural representations for faces and words: A computational exploration, *Cognitive Neuropsychology*, 28:3-4, 251-275

To link to this article: <http://dx.doi.org/10.1080/02643294.2011.609812>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# Complementary neural representations for faces and words: A computational exploration

David C. Plaut and Marlene Behrmann

Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA

A key issue that continues to generate controversy concerns the nature of the psychological, computational, and neural mechanisms that support the visual recognition of objects such as faces and words. While some researchers claim that visual recognition is accomplished by category-specific modules dedicated to processing distinct object classes, other researchers have argued for a more distributed system with only partially specialized cortical regions. Considerable evidence from both functional neuroimaging and neuropsychology would seem to favour the modular view, and yet close examination of those data reveals rather graded patterns of specialization that support a more distributed account. This paper explores a theoretical middle ground in which the functional specialization of brain regions arises from general principles and constraints on neural representation and learning that operate throughout cortex but that nonetheless have distinct implications for different classes of stimuli. The account is supported by a computational simulation, in the form of an artificial neural network, that illustrates how cooperative and competitive interactions in the formation of neural representations for faces and words account for both their shared and distinctive properties. We set out a series of empirical predictions, which are also examined, and consider the further implications of this account.

*Keywords:* Prosopagnosia; Alexia; Neural substrate.

Two opposing theoretical perspectives have been offered to explain the manner by which biological structures, such as the ventral visual cortical regions, come to be functionally optimized for visual object recognition. The first approach argues

that there are distinct cortical modules or subsystems, which mediate particular behavioural processes, such as face, word, and object recognition, in a domain-specific manner (for recent reviews, see Kanwisher, 2010; McKone & Robbins, 2011).<sup>1</sup> Consistent

---

Correspondence should be addressed to David C. Plaut, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213–3890, USA. (E-mail: plaut@cmu.edu).

This work was funded by National Science Foundation (NSF) Grant BCS0923763 to Behrmann and Plaut and by NSF Grant SBE-0542013 to the Temporal Dynamics of Learning Center, an NSF Science of Learning Center. We thank Jennifer Brace for her work in creating the stimuli used in the reported simulation.

---

<sup>1</sup> We use the terms “module” and “modular” not in the strict senses in which Fodor (1983) defined them, but to denote a general class of theoretical commitments in which domain-specific cognitive processes, such as face recognition, are each carried out by a neuroanatomically identifiable cortical area, such as the FFA. To the extent that multiple cortical areas are involved in a given cognitive process, it would mitigate against a modular account of that process but might still be consistent with modular accounts of localized subprocesses.

with this approach is the finding that different regions in extrastriate visual cortex respond selectively to domain-specific categories of visual stimuli: Many recent functional neuroimaging studies have shown, for example, that the fusiform face area (FFA) is activated in response to faces (e.g., Kanwisher, McDermott, & Chun, 1997; Puce, Allison, Gore, & McCarthy, 1995), the parahippocampal place area (PPA) to scenes (e.g., Epstein, Harris, Stanley, & Kanwisher, 1999; Epstein & Kanwisher, 1998; Swards, 2011), the extrastriate body area (EBA) and fusiform body area (FBA) to human bodies and body parts (e.g., Downing, Jiang, Shuman, & Kanwisher, 2001; Peelen & Downing, 2005; Schwarzlose, Baker, & Kanwisher, 2005; Taylor, Wiggett, & Downing, 2010; Willems, Peelen, & Hagoort, 2010), and the visual word form area to orthographic inputs (e.g., Dehaene & Cohen, 2011). Indeed, in each of these regions, the cortical response for the preferred category is about twice that for the nonpreferred category, and this category selectivity can be consistently observed in most normal individuals, even across a range of very different experimental paradigms. All of this attests to the robustness of the evidence that these regions are specialized for, and perhaps even dedicated to, the recognition of particular object classes (Kanwisher, 2010; McKone & Robbins, 2011).

The second approach recognizes the apparent selectivity of neural systems for certain visual classes but argues that this selectivity need not implicate very specialized or dedicated modules per se. This theoretical account entails one or both of two possible brain-behaviour organizations: rather than a single region alone subserving processing of a particular input type (e.g., faces), multiple regions mediate the recognition of a particular object type, and/or an individual region mediates the neural representations of multiple object types. The claim, then, is that, under either of these scenarios, specialization is more graded, and regions may be optimized for, but not necessarily dedicated to, a particular cognitive function. Consistent with this alternative perspective, in addition to the FFA, multiple other cortical regions evince face selectivity, including the

occipital face area (OFA; Gauthier et al., 2000), the posterior superior temporal sulcus (Hoffman & Haxby, 2000), and the anterior temporal lobe (Kriegeskorte, Formisano, Sorge, & Goebel, 2007; Rajimehr, Young, & Tootell, 2009), and, indeed, multiple regions have sufficient neural information to discriminate between individual face exemplars (Nestor, Plaut, & Behrmann, 2011; for more extended review, see Avidan & Behrmann, 2009; Haxby, Petit, Ungerleider, & Courtney, 2000; Ishai, 2008). Furthermore, it is not simply that the distributed network is domain-specific as there are now many functional magnetic resonance imaging (fMRI) studies showing that even highly selective single regions, such as the FFA, evince a blood-oxygen-level-dependent (BOLD) response to different object classes, albeit with lesser degrees of activation than, for example, to faces (e.g., Grill-Spector, Sayres, & Ress, 2006; Hanson & Schmidt, 2011; Haxby et al., 2001; Haxby et al., 2000; Ishai, Schmidt, & Boesiger, 2005; Nestor et al., 2011; Norman, Polyn, Detre, & Haxby, 2006), and the same is true for the visual word form area (VWFA; Nestor, Behrmann, & Plaut, 2011; Price & Devlin, 2011).

In this paper, we compare and contrast the more modular and more distributed accounts with specific reference to two visual classes—faces and words. We choose these two classes not only because, intuitively, they appear to be diametrically opposed but also because they differ obviously along many other dimensions. Words and faces share little in common in their overt geometry, and so their image statistics share minimal, if any, overlap. Additionally, whereas face representations are acquired naturally over the course of experience, word recognition typically requires explicit instruction. Also, whereas faces are probably the most ecologically relevant visual stimuli, orthographies have only been around for a few thousand years, and so the evolutionary trajectories of these two visual classes differ greatly.

We start by reviewing the clear evidence for the separability of the underlying systems for words and faces. Thereafter, we present a proposal in which we argue that common principles may

account for both the similarities and differences in the mechanisms underlying words and faces. We support this proposal with a computational simulation in which a common underlying mechanism, constrained by a putative set of computational principles, mediates both face and word recognition and demonstrates the types of functional specialization observed empirically. Although we address the correspondences between brain and behaviour in these two particular domains, the argument has applicability to other aspects of cognition and its neural correlates, as well, provided that these other cognitive behaviours place the same computational demands on the visual recognition system. We also return to this point in the final discussion.

### Evidence for separability of word and face processing systems

On a modular account of brain-behaviour organization, words and faces engage separate psychological and neural mechanisms and are, essentially, unrelated and independent. Support for this view is substantial and is gleaned from functional imaging investigations, as well as from neuropsychological studies (Kleinschmidt & Cohen, 2006).

#### *The visual word form area*

Numerous functional imaging studies have demonstrated that the word module or “visual word form area” (VWFA; e.g., Cohen et al., 2000; Cohen et al., 2003; Dehaene & Cohen, 2011; Dehaene, Cohen, Sigman, & Vinckier, 2005) responds selectively to visually presented words and letter strings (e.g., Fiez, Balota, Raichle, & Petersen, 1999; Mechelli, Gorno-Tempini, & Price, 2003; Petersen & Fiez, 1993; Petersen, Fox, Snyder, & Raichle, 1990; Turkeltaub, Eden, Jones, & Zeffiro, 2002) to a greater degree than to digits (Polk et al., 2002) or pseudoleters (Allison, McCarthy, Nobre, Puce, & Belger, 1994; Cohen & Dehaene, 2004), but not to spoken words (Cohen & Dehaene, 2004). The VWFA activation is located in left extrastriate cortex (Talairach

coordinates:  $x = -43$ ,  $y = -54$ ,  $z = -12$ ), is identifiable in single subjects (Puce, Allison, Asgari, Gore, & McCarthy, 1996), and is sensitive to the individual’s experience—Hebrew readers show greater activation of this region for Hebrew than for English words and vice versa (Baker et al., 2007), and activation in this area is correlated with literacy (Dehaene & Cohen, 2007, 2011; Dehaene et al., 2010). The VWFA is situated anterior to retinotopic cortex, and, consistent with this, activation is relatively insensitive to retinal position and to the font, size, or case of the input (Polk & Farah, 2002). Activation of VWFA, as measured by event-related potentials (ERPs), is rapid, emerging around 150–200 ms after stimulus onset (McCandliss, Cohen, & Dehaene, 2003). In normal readers, the minimal increase in reaction time (RT) as a function of word length (Lavidor, Ellis, Shillcock, & Bland, 2001; Weekes, 1997) is attributed to the parallel processing of multiple letters (to the limits of foveal acuity, i.e., around 9 letters), and this parallel processing is ascribed to the functionality of the VWFA.

Further support for the circumscribed functionality of the VWFA comes from studies of pre-morbidly literate individuals with “pure alexia” (for review of cases, see Montant & Behrmann, 2000; Starrfelt & Behrmann, 2011). The lesion site in these cases is typically in the left occipitotemporal area along the fusiform and adjacent lingual gyrus, with possible incursion to the inferior longitudinal fasciculus (Cohen, Henry et al., 2004; Cohen et al., 2003; Feinberg, Schindler, Ochoa, Kwan, & Farah, 1994; Salvan et al., 2004) and overlaps the region of the VWFA activation reported above (Hasson, Levy, Behrmann, Hender, & Malach, 2002; Petersen et al., 1990; Puce et al., 1996). The characteristic profile of pure alexia is a linear increase in RT as a function of the number of letters in the input (giving rise to the label “letter-by-letter reading”), and this is assumed to reflect the breakdown of parallel processing in the VWFA and the subsequent reliance on a serial, laborious left-right letter spelling strategy (McCandliss et al., 2003; Warrington & Shallice, 1980). The patients are not aphasic,

typically showing intact production and comprehension of spoken language along with normal writing, all of which supports the circumscribed nature of the problem as a specific difficulty in processing visual word forms (but see Starrfelt & Behrmann, 2011, for discussion of high association with an impairment of number processing as well).

### *The fusiform face area*

Just as in the case of the VWFA, there is substantial evidence for face-processing specificity gleaned from fMRI studies and from patient studies. Functional imaging studies have provided evidence that the region that is functionally specialized for faces, the “fusiform face area” (FFA;  $x = 40, y = -55, z = -10$ ), is selectively activated by faces, especially upright faces, over other nonface objects (Kanwisher, 2010; Puce et al., 1995; Yovel & Kanwisher, 2005) and over animal or cartoon faces (e.g., Kanwisher, 2000; Kanwisher, McDermott, et al., 1997; Kanwisher, Woods, Iacoboni, & Mazziotta, 1997; Sergent, Ohta, & MacDonald, 1992; Spiridon, Fischl, & Kanwisher, 2006), and the magnitude of the activation is correlated with face identification ability (Furl, Garrido, Dolan, Driver, & Duchaine, 2010; Yovel, Tambini, & Brandman, 2008). FFA activation is situated anterior to retinotopic cortex, and, consistent with this, activation is relatively insensitive to retinal position and to size, colour, format (drawing or photographs), and viewpoint of input. The FFA is selectively activated for faces but abuts other cortical regions that are specialized for other visual categories, such as scenes, animals, and tools (e.g., Reddy & Kanwisher, 2006; Spiridon & Kanwisher, 2002; also, Puce et al., 1996; Puce et al., 1995; Tranel, Damasio, & Damasio, 1997).

Correspondingly, lesions to the FFA (Bouvier & Engel, 2006; Damasio, Damasio, & Tranel, 1986; Kleinschmidt & Cohen, 2006) result in prosopagnosia, a selective impairment in face recognition. The lesion in prosopagnosia is often bilateral, affecting the temporo-occipital cortex in the region of the FFA, but unilateral right-hemisphere lesions to this same region may

suffice to give rise to this disorder (Barton, 2008; Bouvier & Engel, 2006), and prosopagnosia can also be congenital or developmental in the absence of a frank lesion (Behrmann & Avidan, 2005). The difficulty in recognizing faces can be dramatic, including failures to recognize friends or even close family members. Unlike normal observers, these individuals do not obviously exhibit the advantage for upright over inverted faces (occasionally even showing an inversion superiority effect; Farah, 1996; Farah, Tanaka, & Drain, 1995) and do not appear to process faces configurally, thus failing to evince the benefit from the presence of the whole face over just parts of the face (Barton, 2009; Barton, Cherkasova, Press, Intriligator, & O'Connor, 2004; Busigny & Rossion, 2010; Tanaka & Farah, 1993).

Taken together, these studies provide empirical support for the claim that there is specialized processing of faces and words associated with two distinct cortical modules, the FFA for faces and the VWFA for words, and that these two systems are separable and independent.

### **Not only differences but also commonalities**

Although there is general consensus that the FFA and VWFA are tuned to faces and words, respectively, there are also intriguing empirical data that suggest that both their tuning and their hemispheric specialization is relative or graded. For example, it appears that both the VWFA and FFA can be activated by a wide range of stimuli, not just faces or words: The VWFA is strongly activated in response to chequerboards, pictured objects, and verb naming to pictures (Devlin, Jamison, Gonnerman, & Matthews, 2006; Murtha, Chertkow, Beauregard, & Evans, 1999; Price & Devlin, 2003, 2011) and even to nonvisual inputs such as Braille (Büchel, Price, & Friston, 1998; Reich, Szwed, Cohen, & Amedi, 2011), whereas the FFA is activated by a range of nonface stimuli, such as houses and cars, but also novel objects such as Greebles (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999) and chess configurations (Bilalic, Langner, Ulrich, &

Grodd, 2011), although the full extent of the selectivity is still controversial (Gauthier et al., 1999; Grill-Spector et al., 2006; Haxby, 2006). Neither the FFA nor the VWFA, however, appears to be as strongly activated by these other stimuli as is the case when shown the “preferred” input type, reflecting perhaps the graded nature of the underlying representations. Recent imaging studies adopting multivariate methods applied to fMRI data of ventral visual cortex have begun to uncover the co-mingling of patterns of activation associated with different stimulus types (e.g., face and word representations) to an even greater degree than was revealed in earlier studies employing univariate analyses (for an example of a recent study using multivoxel pattern analysis, see Nestor et al., 2011). In these multivariate studies, it is not simply the magnitude of the activation that is crucial but the distribution of the neural information in the patterns of voxel activation.

Also relevant to the similarities across classes is the observation that almost all fMRI and ERP studies show bilateral activation for words and for faces, albeit with differential hemispheric asymmetry and greater scalp potential for the preferred stimulus type in the corresponding hemisphere—words on the left and faces on the right (e.g., see Hasson et al., 2002; Kanwisher, McDermott & Chun, 1997; Kronbichler et al., 2004; Price & Mechelli, 2005; Puce et al., 1996; Sergent et al., 1992; Tagamets, Novick, Chalmers, & Friedman, 2000). Moreover, the peak activation for words in VWFA and for faces in FFA (although coordinates differ a little across different studies) are very comparable in the two hemispheres (for example, Talairach coordinates for peak for words,  $x = -43$ ,  $y = -54$ ,  $z = -12$ , and for faces,  $x = 40$ ,  $y = -55$ ,  $z = -10$ ). We also note that these coordinates roughly demarcate cortical sites that are anterior to retinotopic cortex but are situated in what would be the anterior extrapolation of the fovea (Hasson et al., 2002; Levy, Hasson, Hendler, & Malach, 2001). The localization of the functional regions in this cortical location is consistent with the invariance of face and word activation over

retinal position of the inputs, but also with the fact that reliance on fine-grained visual discrimination is a necessary component of both face and word recognition.

Somewhat surprisingly, there has not been a systematic examination of the word recognition of prosopagnosic individuals and the face recognition of pure alexic individuals. There are some hints, however, that each hemisphere may play a dual, albeit graded, role in both face and word recognition. For example, it has been reported that the face recognition impairment is more severe following bilateral than unilateral lesions (Damasio et al., 1986; Gainotti & Marra, 2011), implicating both hemispheres to some extent, and that transcranial magnetic stimulation (TMS) of the right hemisphere (RH) even impairs reading in patients with left-hemisphere (LH) lesions (Coslett & Monsul, 1994).

Additionally, in a few case studies in which both stimulus classes have been examined, some prosopagnosic individuals show increased word length effects in reading aloud single words: For example, the slope of the reaction time in single word reading was 104 ms and 241 ms per additional letter for prosopagnosic patients S.M. and R.N., respectively, compared with the normal slope of about 10 ms for words 3 through 8 letters in length (Behrmann & Kimchi, 2003). In a complementary fashion, there have only been a few reports of pure alexic individuals who have difficulties with face recognition (also see Farah, 1991, 1992, 1999, for listing of co-occurrences of different forms of agnosia/alexia), although this is not always assessed in these cases. One recent relevant study documents a case with left occipital arteriovenous malformation in whom both pure alexia and prosopagnosia were evident (Liu, Wang, & Yen., 2011). Many studies do report abnormalities in the recognition of non-orthographic stimuli in pure alexia even after a unilateral lesion (Behrmann, Nelson, & Sekuler, 1998; Starrfelt & Behrmann, 2011), and so one might predict a decrement in face recognition in these cases, as well. As evident, closer scrutiny of the existing data, to the extent they are available, suggests that there may be more overlap in face

and word processing in the preeminent face (right FFA) and word (left VWFA) regions than originally considered. To account for both the apparent differences and the similarities, we propose an account that differs from the strictly modular or domain-specific view.

### **An alternative proposal: Common constraints on faces and words**

The theoretical proposal outlined in this paper adopts an alternative perspective with respect to the key systems engaged in face and word processing. The central idea is that visual object recognition (e.g., face and word recognition) is supported, not by highly specialized (or dedicated) modules per se, but by a distributed and interactive network of brain regions with similar computations but whose organization is strongly shaped and modified by experience. This view then incorporates both the claim that multiple cortical regions are engaged and that the nodes of this distributed network play a role in representing more than one stimulus type. Importantly, on this view, the functional specialization of brain regions is graded rather than absolute and reflects the consequences of a set of general principles and constraints on neural computation that operate throughout cortex but that nonetheless have distinct implications for different classes of stimuli. Note that, on this account, there is no appeal to prespecified modules, and, rather than claiming de facto sensitivity to different visual classes, the origin and emergence of these graded mechanisms is captured too. The novelty of this approach is not the principles themselves (see both Dehaene & Cohen, 2011, and Price & Devlin, 2011, for similar notions about the VWFA) but their integrated application to derive common consequences for cortical organization and behaviour for words and faces.

This alternative proposal takes as its starting assumptions three general principles of neural computation: that the neural system for face/word recognition is distributed, that knowledge is represented in this system by cooperation and competition between the processing units, and

that the organization of the system is constrained by topographical considerations, pressure for proximity, and the division of labour between the two hemispheres of the brain. We expand on these assumptions here.

### *Distributed representation and knowledge*

We assume that the neural system for visual object recognition consists of a set of hierarchically organized cortical areas, ranging from local retinotopic information in V1 through more global, object-based, and semantic information in anterior temporal cortex (Grill-Spector & Malach, 2004). At each level, the visual stimulus is represented by the activity of a large number of neurons, and each neuron participates in coding a large number of stimuli. Generally, stimuli that are similar with respect to the information coded by a particular region evoke similar (overlapping) patterns of activity. The set of constraints on how activity at one level produces activity at the next level—that is, the knowledge of how features combine to form features at the next level—is encoded by the pattern of synaptic connections and strengths between and within the regions. Learning involves modifying these synapses in a way that alters the representations evoked by visual stimuli—typically in a way that captures the relevant information in the domain better and that supports more effective behavioural outcomes. With extended experience, expertise develops through the refinement, specialization, and elaboration of representations, requiring the recruitment of additional neurons and a larger region of cortex (Quartz & Sejnowski, 1997).

### *Representational cooperation and competition*

As illustrated by artificial neural networks (e.g., McClelland & Rumelhart, 1985), a single pattern of synaptic connections can learn to encode the knowledge needed to represent many stimuli, but its ability to do so depends on the degree to which the relevant knowledge is consistent or systematic (i.e., similar representations at one level correspond to similar representations at another). In general, systematic domains benefit from highly overlapping neural representations

that support generalization, whereas unsystematic, unrelated domains require largely nonoverlapping representations to avoid interference. Thus, if a cortical region represents one type of information, it is ill-suited to represent another type of information that requires unrelated knowledge, and so that information must be represented by a different region. On the other hand, effective cognitive processing requires the coordination of multiple levels of representation within a given domain, and often across multiple domains. Of course, representations can cooperate directly only to the extent they are connected—that is, there are synapses between the regions encoding the relevant knowledge of how they are related; otherwise, they must cooperate indirectly through mediating representations. In this way, the neural organization of cognitive processing is strongly constrained by available connectivity (see Mahon & Caramazza, 2011, for a similar argument regarding connectivity serving as an endogenous constraint on topographic organization in the ventral stream).

#### *Topography, proximity, and hemispheric organization*

Brain organization must permit sufficient connectivity among neurons to carry out the necessary information processing, but the total axonal volume must fit within the confines of the skull (for similar discussion, see Cowey, 1979). This constraint is severe: If the brain's  $10^{11}$  neurons were placed on a sphere and were fully interconnected with 0.1-mm radius axons, accommodating the axon volume would require a sphere over 20 km in diameter (Nelson & Bower, 1990). If we think of brain organization as the result of a complex optimization process that minimizes “costs” associated with the degree to which various pressures or biases are violated, then clearly there is a strong pressure to keep connectivity as local as possible. Long-distance projections are certainly present in the brain but they are relatively rare and presumably play a sufficiently critical functional role to offset their cost in volume. In fact, the organization of human neocortex as a folded sheet can be understood as a

compromise between the spherical shape that would minimize long-distance axon length and the need for greater cortical area to support highly elaborated representations. The organization into two hemispheres is also relevant here, as interhemispheric connectivity is largely restricted to homologous areas and is thus vastly less dense than connectivity within each hemisphere. Even at a local scale, the volume of connectivity within an area can be minimized by adopting a topographic organization so that related information is represented in as close proximity as possible (Jacobs & Jordan, 1992). This is seen most clearly in the retinotopic organization of early visual areas, given that light falling on adjacent patches of the retina is highly likely to contain related information. Note that the dimensions of this topography are not in the Cartesian ( $x, y$ ) coordinates that apply naturally to images, but something closer to polar ( $r, \theta$ ) coordinates, where eccentricity (central vs. peripheral) is coded along one axis, and rotational angle is coded along another (e.g., De Yoe et al., 1996; Grill-Spector & Malach, 2004; Sereno et al., 1995; Tootell et al., 1997). The relevant dimensions of similarity for higher level visual areas are, of course, far less well understood, but the local connectivity constraint is no less pertinent (Jacobs, 1997).

Despite these commonalities, the principles rule out using the very same cortical region to represent both faces and words because these stimuli require entirely distinct primitives to be represented as visual objects and typically have distinct consequences for cognition (faces designate individuals, whereas—apart from proper names—words designate objects, actions, properties, typically at a basic rather than individual level). Given the need for written words to interact with aspects of language that are left lateralized in most individuals (Cai, Lavidor, Brysbaert, Paulignan, & Nazir, 2008), it follows from representational competition and cooperation that visual word representations would be predominantly located in the dominant language hemisphere (Price & Devlin, 2011) whereas face representations would be located in the



homologous region in the RH. Indeed, some data to support this competition/cooperation come from the observation that, with increasing literacy, there is a decrease in response to faces in the VWFA (Dehaene et al., 2010) and that, in four-year-olds, performance in identifying alphanumeric characters (digits and letters) is correlated with a decrease in left fusiform activity (Cantlon, Pineda, Dehaene, & Pelphrey, 2011). In both of these studies, however, the competition appears to be restricted to the left hemisphere, with the trading relations between faces and words manifesting in the left fusiform region. The hypothesis we propose, however, encompasses both the left and right hemispheres, with the competition and cooperation playing out for words and faces across both sides of cortex.

## SIMULATION

To date, the majority of computational work on face recognition has an applied focus with only tangential relevance to the human cognitive and neural system—this includes approaches based on principal components analysis (e.g., Turk & Pentland, 1991), independent components analysis (e.g., Bartlett, Movellan, & Sejnowski, 2002), linear discriminant analysis (e.g., Etemad & Chellappa, 1997), kernel methods (e.g., Bach & Jordan, 2002; Yang, 2002), 3D morphable models (e.g., Blanz & Vetter, 2003), and Bayesian inference (e.g., Moghaddam, Jebara, & Pentland, 2000). Modelling efforts that explicitly address psychological and neuropsychological issues (e.g., Burton, Young, Bruce, Johnston, & Ellis, 1991; Farah, O'Reilly, & Vecera, 1993) have tended to focus on the interaction of higher level knowledge with rather less consideration of low- and intermediate-level visual representation and processing (although see Burton et al., 1999). More recently, Cottrell and colleagues (Dailey & Cottrell, 1999; Dailey, Cottrell, Padgett, & Adolphs, 2002; Hsiao, Shieh, & Cottrell, 2008; Kanan & Cottrell, 2010; Tong, Joyce, & Cottrell, 2008) have extended this work by coupling distributed network modelling with

more realistic assumptions about early visual processing.

A similar situation holds with regard to word recognition. Although some early cognitive and neuropsychological modelling employed hierarchical visual representations of letters and words (McClelland & Rumelhart, 1981; Mozer, 1991; Mozer & Behrmann, 1990), the vast majority of more recent work has emphasized higher level interactions of orthographic, phonological, and semantic knowledge (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Harm & Seidenberg, 1999, 2004; Perry, Ziegler, & Zorzi, 2007; Plaut, McClelland, Seidenberg, & Patterson, 1996; although see Plaut, 1999). Efforts to model orthographic representations per se (e.g., SOLAR, Davis, 1999; SERIOL, Whitney, 2001) have typically focused more narrowly on letter position effects in orthographic priming. One notable exception is the split-fovea model (Shillcock, Ellison, & Monaghan, 2000), which explicitly considers the representational implications of a divided visual field. Although this specific model runs into some empirical difficulties (see, e.g., Grainger, Granier, Farioli, Van Assche, & van Heuven, 2006), there is no doubt that the cortical representation of words is shaped in important ways by hemispheric organization and specialization (Cai et al., 2008).

Given the apparent lack (to date) of any proposed relationship between face and word processing, it is not surprising that our computational work is the first to address these domains together within a single model. Although the current implementation does not extend to the higher level knowledge involved in face and word recognition, the underlying principles are fully compatible with ongoing modelling work at these higher levels.

Perhaps the least familiar of our computational principles concerns the impact of local connectivity on learning. Thus, as an initial exploration of the impact of topographically constrained learning on cortical organization, we carried out a simulation in which an artificial neural network was trained to take retinotopic visual information as input and to map this via hemisphere-specific

intermediate representations (corresponding to left and right occipitotemporal cortex) to recognize faces, words, and—as a commonly used contrasting category—houses. The topographic bias on learning, combined with the demands for high-acuity information for face and word recognition, should lead to these stimuli being represented in intermediate (fusiform) regions near central vision. The need for representational cooperation between words and language-related information, in conjunction with representational competition between faces and words (given their incompatibility as visual objects), is predicted to give rise to left-hemisphere specialization for words and right-hemisphere specialization for faces. Due to the graded nature of the learning constraints, this specialization should be only partial, with both regions participating in processing both types of stimuli to a certain degree. Houses are expected to be represented by more peripheral regions in fusiform cortex, analogous to the “parahippocampal place area” (PPA; Epstein & Kanwisher, 1998; Levy et al., 2001).

All simulations were developed within the Light Efficient Network Simulator (Lens; Version 2.63), developed by Doug Rohde and available for download at <http://tedlab.mit.edu/~dr/Lens>.

## Method

### *Stimuli*

As the goals of the current work are to explore and illustrate the implications of a set of putative computational principles rather than to build a realistic model of visual face and word perception, the task and network architecture employed in the simulation were kept as simple as possible. The stimuli used in the simulation were derived from  $32 \times 32$ -bit schematic line drawings of faces, houses, and words that embodied critical differences in the demands of recognition of these classes of stimuli (see Figure 1). Each of 34 faces differed in terms of small changes in the positions or shapes of central features (e.g., separation and height of eyes, length of nose, height and width of mouth). Each of 40 three-letter (CVC, where C = consonant, V = vowel) words were created

from combinations of five possible letters for each position and, like faces, differed from each other only in terms of features within central vision. By contrast, each of nine houses differed in terms of properties that varied across the entire visual field (e.g., size of windows, number of eaves, presence of porch, size of base). Note that these rather small differences place high demands on fine visual acuity to ensure accurate discrimination between exemplars.

Each item was presented at nine different scales, ranging from 1.0 to 0.6 in steps of 0.05, for a total of 747 input patterns. For each pattern, retinotopic input activation was generated by smoothing the original bit patterns by convolving them with a Gaussian ( $SD = 0.5$ ) and then transforming the resulting values into polar coordinates ( $r, \theta$ ), such that eccentricity  $r$  varied along the horizontal axis (with central information on the left and peripheral information on the right), and visual angle  $\theta$  varied along the vertical axis (see Figure 2 for examples).

As the current work is concerned only with the nature of the visual representations of various stimulus classes, no attempt was made to approximate the structure of higher level information that such representations provide access to, beyond the need to identify (individuate) each unique face, word, and house (despite changes in scale). Accordingly, the output representations used in the simulation consisted of individual “localist” units for each of the 34 faces, 40 words, and 9 houses. We recognize, of course, that this is implausible as the actual output of the human system.

### *Network architecture*

The network architecture is depicted in Figure 3. In the model,  $32 \times 32$  retinotopic visual input to each hemisphere (in polar coordinates) is mapped via 64 ( $32 \times 2$ ) intermediate units in each hemisphere (corresponding to fusiform cortex) onto a set of 83 “identity” units (one for each unique word, face, and house). In addition, to approximate the influence of a left-hemisphere specialization for language, word inputs were also trained to activate one of a set of 40 “language” units

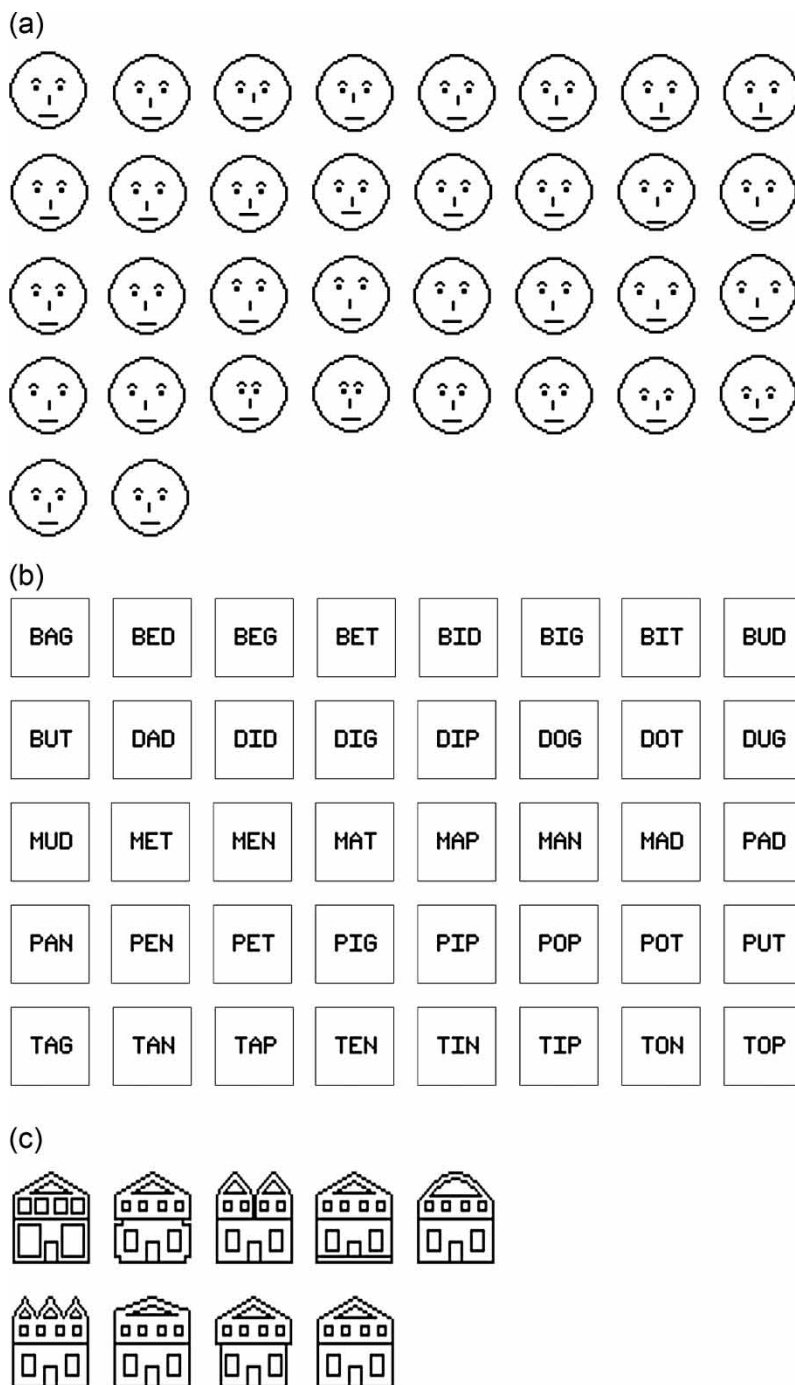
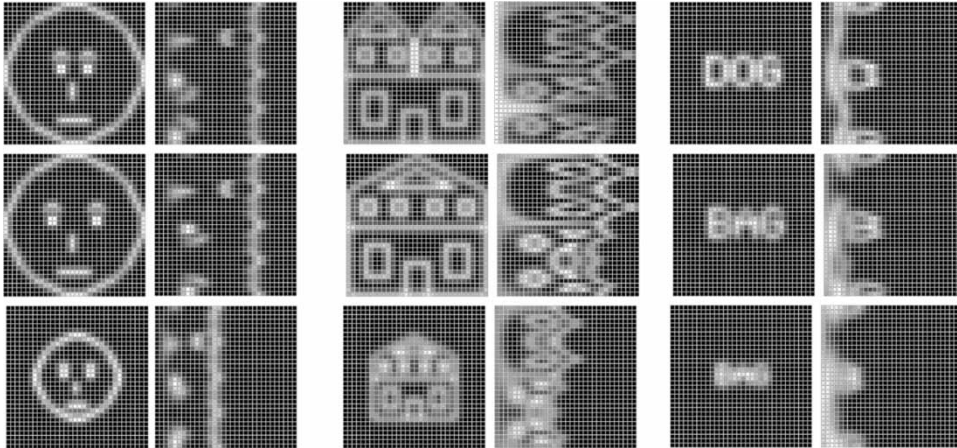
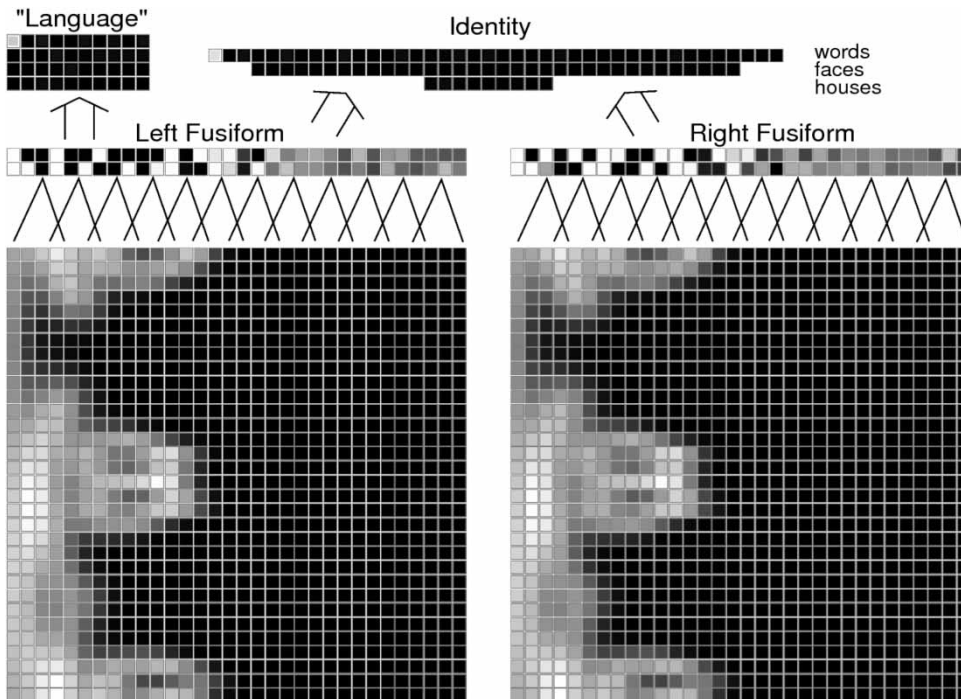


Figure 1. The full set of face, word, and house stimuli used to create inputs to the simulation. Each picture defines a unique identity; the actual inputs to the network were generated by smoothing and transforming into polar coordinates (see Figure 2 for examples).



**Figure 2.** Example face, house, and word stimuli after Gaussian smoothing. For each stimulus class, the left three panels show stimuli in x-y coordinates; the corresponding right panel shows the same stimulus in polar coordinates—the form actually presented to the network as input. Also, for each class, the top two rows differ in identity; the bottom two rows differ only in scale.



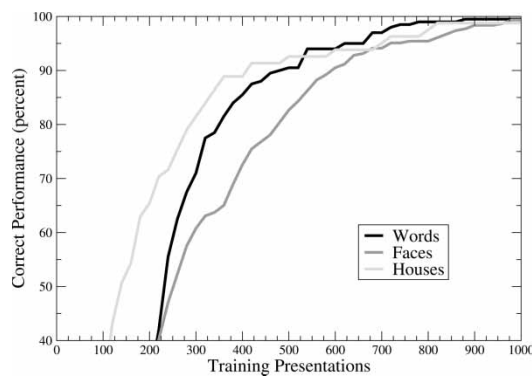
**Figure 3.** The network architecture used in the simulation. Each square corresponds to a particular processing unit. Activations for a particular example input are depicted by the greyscale value of the square (black = 0.0; white = 1.0). Sets of connections are depicted by lines/arrows but are not shown in detail. For each of the two retinotopic input layers, activations toward the left encode central visual information, whereas activations toward the right encode more peripheral information. The projections from the two input layers to the left and right intermediate (fusiform) layers are subject to a horizontal topographic bias favouring short connections.

that receive input solely from the left-hemisphere intermediate units. Finally, the input-to-intermediate connections were subject to strong topographic bias during learning. Although this bias is assumed to be enforced in the brain by the relative density of synapses as a function of distance, the small scale of the simulation made it more appropriate to implement this bias slightly differently. Specifically, the input units were fully connected to the intermediate units, but the efficacy of learning decreased as a Gaussian function ( $SD = 3.0$ ) of the distance between the connected units (Plaut, 2002). As we were primarily concerned with the impact of eccentricity on learning, this metric considered only horizontal distance in the simulation (i.e., all units in the same column in Figure 2 were considered to have the equivalent functional position). Thus, in practical terms, learning was effective on connections from inputs directly “below” a given intermediate unit, but increasingly ineffective on connections from units with progressively different horizontal positions. For similar reasons, although we would claim that all connectivity in the brain is subject to a topographic bias, we did not apply this bias to any of the intermediate-to-output connections because we had no hypothesis concerning the relative proximity of semantic or identity information (beyond the left-lateralization of language information). The important consequence of this is that the identity units for faces, words, and

houses are equivalent in their connectivity with the intermediate units, and thus any distinction in the specialization of the intermediate units must arise solely from properties of the inputs. Finally, the simulation employed a feedforward architecture, without lateral (within-layer) or top-down connections, solely for computational convenience.

### *Training and testing*

When presented with a scaled version of each face, word, or house, the network was trained to activate the correct identity unit (and, for words, the correct “language” unit). Back-propagation (Rumelhart, Hinton, & Williams, 1986) was used to calculate how to change each connection weight in the network to reduce the discrepancy between the output activation pattern generated by the network and the correct pattern. Although not biologically plausible in literal form, back-propagation is functionally equivalent to more plausible procedures such as contrastive Hebbian learning (see, e.g., O’Reilly, 1996). The topographic bias on learning at the intermediate layer was implemented by scaling these weight changes by a decreasing (Gaussian) function of the horizontal distance between the connected input and intermediate units. Following training, the network was considered to be correct if, for a given input, the correct identity unit was more active than any other.



**Figure 4.** Correct performance of the network in identifying faces, words, and houses as a function of the number of presentations of the entire set of 747 example patterns received by the network during training.

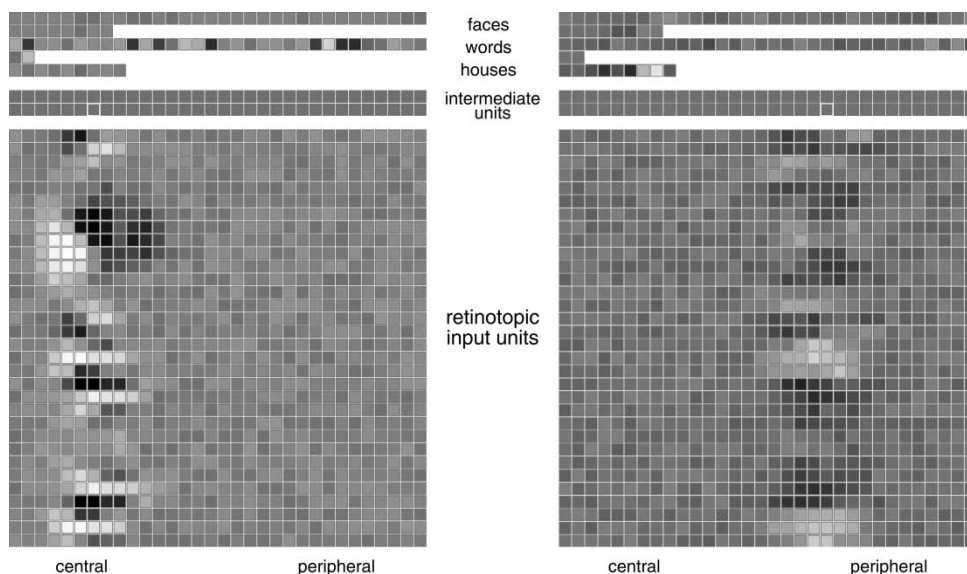
The primary basis for establishing specialization in the network was its performance following lesions restricted to spatially contiguous areas of either the left or right intermediate (fusiform) layer. Lesions were administered by removing three adjacent columns (6 units in total) from one of these layers and evaluating the performance of the damaged network for each of the 747 input patterns (83 identities  $\times$  9 scales). The horizontal position of these lesions was varied systematically in order to evaluate the relative specialization of each intermediate layer for each stimulus class as a function of visual eccentricity.

### Results and discussion

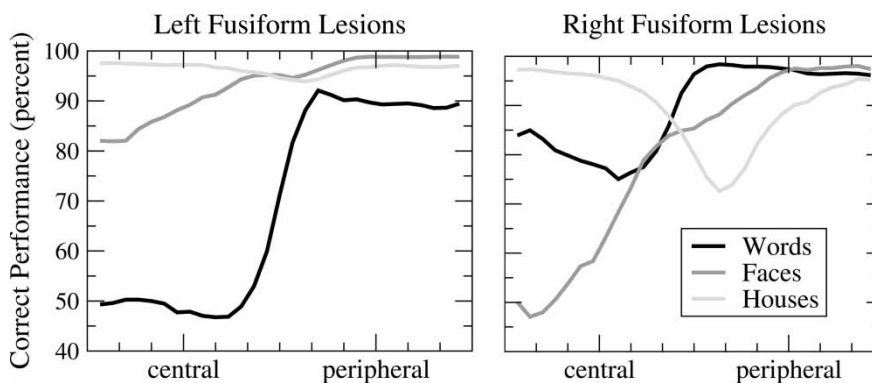
After 1,000 training presentations of each pattern, the network is fully accurate at recognizing instances of each face, word, and house (see Figure 4). Over the course of acquisition, performance on houses is better than on the other stimulus

classes because there are fewer of them to differentiate. Performance on words is better than on faces in part because the latter involve more subtle featural distinctions and, in part, because the extra demands of activating “language” information for words provide additional error (and therefore learning). By the end of training, however, performance on all three classes is equivalent and at ceiling.

To illustrate the effects of the topographic bias on learning in the network, Figure 5 shows examples of the “receptive” and “projective” fields learned by two intermediate units. The left display is for a unit that has a receptive field in central vision (i.e., toward the left of the retinotopic input) and has output that is largely selective for faces (third and fourth rows in the top group of units). By contrast, the right display is for a unit that has a more peripheral receptive field (i.e., toward the centre or right of the input) and is largely selective for houses (last row in the top group). These weight diagrams illustrate



**Figure 5.** Example receptive (*input-to-intermediate*) and projective (*intermediate-to-output*) fields for two units in the right-hemisphere intermediate layer. Each square shows the value of the weight (lighter for positive weights; darker for negative weights) from that unit either into or out of the depicted unit (outlined in the middle layer). The top group of units are output units; the first two rows are for words, the next two rows are for faces, and the last row is for houses.



**Figure 6.** Correct performance on faces, words, and houses following focal lesions to either the left or the right intermediate (fusiform) layer in the network, as a function of the horizontal position of the lesion (ranging from central to peripheral moving left to right within each hemisphere).

the impact of the topographic constraint on learning and provide indirect evidence for learned category specificity of intermediate units as a function of their eccentricity.

More direct evidence for topographic specialization comes from the effects of localized lesions to the network. Figure 6 shows the performance of the network for each category of stimuli as a function of the horizontal position of lesions to three adjacent columns of units in either the left or right intermediate (fusiform) layer. Left-hemisphere lesions to the region of the fusiform near central visual information (analogous to the VWFA) produce a marked impairment in word recognition, but also a milder impairment in face recognition (relative to house recognition). By contrast, lesions to the corresponding region in the right hemisphere (analogous to the FFA) impair face recognition most, but also word recognition to a lesser extent. Finally, lesions to the right fusiform adjacent to more peripheral visual information (analogous to the PPA) produce the greatest impairment on houses and a milder impairment on faces (relative to words). These findings are in qualitative agreement with existing observations (Epstein, Deyoe, Press, Rosen, & Kanwisher,

2001) and our derived predictions for graded functional specialization.<sup>2</sup>

This small-scale simulation provides a simple but clear demonstration of the impact of a topographic constraint on learning on the organization of representations mediating face, word, and house recognition. The reliance of face and word recognition on central, high-acuity information leads to selectivity in the intermediate (fusiform) units closest to this information (Hasson et al., 2002; Levy et al., 2001; Levy, Hasson, & Malach, 2004). Competition between inconsistent information (and cooperation between word representations and language-related information) leads to substantial but still graded hemispheric specialization, with words represented primarily on the left and faces primarily on the right.

## GENERAL DISCUSSION

This paper takes at its starting point a debate about the manner in which the brain is organized in the service of behaviour. One longstanding view is that different parts of the brain are specialized for, and perhaps even dedicated to, different cognitive

<sup>2</sup> Although not reported here in detail, these qualitative results are stable over changes to nonessential aspects of the network architecture and training methods, including variations (within reasonable limits) in random initial weights, learning parameters, and numbers of hidden units.

functions. An alternative account is one in which there is no unique, one-function one-region correspondence; rather, a single region subserves many different tasks, and/or a single task is mediated by many different regions. We have explored this latter, distributed perspective in the context of the ventral visual cortex and its organizational structure, taking as a model the case of face and word recognition. This is a particularly good domain in which to explore these different theoretical accounts as there are considerable empirical data on both the psychological and neural mechanisms involved in these functions, derived primarily from neuroimaging and neuropsychological investigations. We note, however, that the theoretical proposal is more general and applies to other visual domains that have the same computational demands as do words and faces.

The computational account we propose puts forward the theoretical claim that the representations of faces and words, albeit so apparently different in their surface characteristics and their underlying neural substrate, are the product of the same computational principles. Specifically, the visual recognition system has at its core three general principles, all of which have profound implications for both the commonalities and differences in the neural organization and functionality of face/word processing. The key principles are:

1. *Distributed representation and knowledge:* Visual objects are represented by distributed patterns of neural activity within a hierarchically organized system, where learning involves modifying the pattern of synaptic connectivity between neurons within and between regions on the basis of experience.
2. *Representational cooperation and competition:* Effective cognitive processing requires the coordination of related information across multiple levels of representation, whereas unrelated or incompatible information must be represented over separate regions to avoid interference.
3. *Topography, proximity, and hemispheric organization:* Representational cooperation must be

accomplished with largely local connectivity between topographically organized brain regions, and with limited connectivity between hemispheres, so that total axonal volume fits within the confines of the skull.

These constraints have a fundamental impact on how faces and words are represented and processed within the visual system.

Here, we show that a small-scale simulation that is trained to recognize faces, words, and houses in a manner consistent with these principles illustrates how a topographic constraint on learning can give rise to learned category specificity of intermediate units as a function of their eccentricity. This topographic constraint is further anchored by the reliance on fine-grained visual processing for discriminating subtle visual differences among words and among faces. Specifically, left-hemisphere lesions to the region of the intermediate layer near central visual information (analogous to the VWFA) produce a marked impairment in word recognition, but also a milder impairment in face recognition (relative to house recognition). By contrast, lesions to the corresponding region in the right hemisphere (analogous to the FFA) impair face recognition most, but also word recognition to a lesser extent. These findings are in qualitative agreement with existing observations and our derived predictions for graded functional specialization.

The results of the simulation provide an existence proof of a system in which face and word recognition are subject to the same computational constraints but in which relative specialization, by virtue of other competing pressures (to restrict connection length specifically with language areas, see Cai et al., 2008; Price & Devlin, 2011) also demonstrates some functional specialization. The idea that there are both many similarities as well as differences among the mechanisms supporting face and word recognition is endorsed by the existing neuroimaging studies (many or even most of which reflect bilateral activation for faces and words) and some neuropsychological studies, which show dual impairment following a unilateral hemispheric lesion.



## Converging evidence

There are some additional considerations that favour a common-mechanism perspective. Both face and word recognition are domains with which most individuals have extensive experience and expertise. Both classes place demands on high-acuity information to encode subtle but critical visual information, and, thus, the fact that the cortical regions that are selective for face and word processing are located adjacent to the central visual information within the highest level of retinotopic representation (Hasson et al., 2002; Levy et al., 2001) can be understood as a natural consequence of topography and the constraint on local connectivity. Furthermore, the cortical regions selective for these stimuli come to be located adjacent to retinotopic regions coding foveal information, but in different hemispheres, with words in the left to permit coordination with other language-related knowledge. Additionally, the VWFA and FFA are both insensitive to low-level input variations (e.g., letter font; viewpoint for faces) suggesting that both regions reflect functional specialization of higher order visual cortex.

A further commonality is that lesions to each region render the individual reliant on a more piecemeal or segmental approach rather than one in which the configural or whole is accessible. Just as VWFA lesions result in a laborious letter-by-letter sequential decoding of the individual letters, lesions resulting in prosopagnosia give rise to a similar sequential process with greater reliance on some features of the face (e.g., mouth, see example in Bukach, Le Grand, Kaiser, Bub, & Tanaka, 2008) and a laborious encoding of the features as reflected in eye movement patterns (e.g., see Stephan & Caine, 2009). Also, just as VWFA activation is affected by experience (Baker et al., 2007; Wong et al., 2005), so adult-like face processing in the FFA also emerges with age and

experience, and both regions evince a protracted developmental trajectory with signs of specificity emerging roughly when children are in elementary school (Brem, Bach et al., 2010; Brem, Bucher et al., 2006; Cohen Kadosh & Johnson, 2007; Cohen Kadosh et al., 2010; Golarai et al., 2007; Golarai, Liberman, Yoon, & Grill-Spector, 2010; Joseph, Gathers & Bhatt, 2011; Scherf, Behrmann, Humphreys, & Luna, 2007).<sup>3</sup> Finally, both regions are rather plastic: The VWFA can be acquired in the RH after left occipital resection in childhood (Cohen, Lehericy, et al., 2004), and there may also be a shift to the right in patients following acquired LH lesions (Cohen, Lehericy, et al., 2004; Cohen et al., 2003). Relatedly, there are no apparent differences in the face recognition deficits of individuals with unilateral lesions in infancy that impacted either the LH or the RH (de Schonen et al., 2005).

One apparent challenge to our emphasis on the role of visual experience in shaping the cortical organization of face and word processing is the recent observation by Reich et al. (2011) that the VWFA, as localized in sighted individuals, is also the location of peak activation in congenitally blind individuals reading Braille words (compared to nonsense Braille control stimuli). It should be noted, however, that Reich et al. found reliable differences in the entire left ventral occipitotemporal cortex all the way to V1, consistent with earlier findings by Noppeney, Friston, and Price (2003) that were attributed to recruitment of these areas for semantic processing. Reich et al.'s only evidence for (tactile) word-form representations per se was that activation differences were larger for Braille reading than for auditory verb generation, but this comparison was possibly dubious as it involved separate and unrelated control tasks. Moreover, the fact that activation differences peaked in the same location in blind and sighted individuals may arise because of intrinsic

<sup>3</sup>The exact nature of the change over developmental time remains somewhat controversial with some studies showing changes in the volume of activation for one category over another and others showing a change in functional/effective connectivity over the course of development. The studies are also not entirely consistent with each other (see Cantlon et al., 2011, showing adult-like activation to faces as well as sensitivity to alphanumeric symbols in four-year-olds although volume/cluster size was not evaluated in this study). These empirical discrepancies remain to be resolved.

patterns of connectivity between early visual cortex and parietal structures involved in spatial attention (Greenberg et al., 2011), and this inherent white matter arrangement biases where activation peaks are likely to be observed in functional imaging studies.

### Where do hemispheric differences come from?

Our account presupposes that the hemispheric differences and asymmetries emerge over the course of experience, during which time the connectivity and topography constraints play out. There are, however, other theoretical stances, which interpret the emergent hemispheric differences as arising from a different origin. One such account is that the left-hemisphere (LH) processes input in an analytical or part-based way (hence its role with words) whereas the right-hemisphere (RH) processes input more holistically (hence its role with faces; see also Farah 1991, 1992). A second view attributes the word/face distinction to the distinction between linguistic (LH) and spatial (RH) processing. Yet a further view is that the face/word difference arises from differential frequency sensitivity (Robertson & Ivry, 2000) with the RH and LH responding to relatively low and high spatial frequency information, respectively, and the former critical for faces (RH), and the latter for words (LH). A final possibility is that the face/word differences arise from the differential predisposition to process inputs categorically (LH words) versus by coordinate relations (RH faces; Kosslyn et al., 1989). These accounts all suggest that it is the fundamental (perhaps even hard-wired or innately specified) properties of the two hemispheres that play a role in shaping the underlying computational differences between words and faces, but how and to what extent this is so remains to be determined.

Our approach is not mutually exclusive with these process-based accounts but it emphasizes the importance of expertise in shaping cortical organization and function (Gauthier & Bukach, 2007; Gauthier & Nelson, 2001), although

expertise alone cannot explain why the FFA and VWFA are located where they are, nor why other types of expertise (namely, those not demanding high-acuity visual information) do not engage these areas. It also shares the fundamental assumption of Malach and colleagues (Levy et al., 2001; Levy et al., 2004) regarding pressure for foveal acuity and cortical topography, but goes beyond this by implementing the ideas in explicit simulations, enabling the testing of specific predictions concerning the relationship of face and word processing. There remain, of course, complex questions about why the left hemisphere is language dominant in the majority of the population and the source of this organizational pattern. Such issues are beyond the scope of this paper but are intriguing and remain to be addressed, too.

### Predictions

A full assessment of the tractability of our account remains to be undertaken, and, in particular, there are a number of predictions that can be tested. Much of our work, thus far, has focused on the need for representational cooperation between words and the language-related output, as the key pressure that drives the left-hemisphere specialization. But this cooperation occurs in conjunction with representational competition between faces and words (given their incompatibility as visual objects), and this competition, too, motivates the hemispheric distinctions with left and right biased for words and faces, respectively. A prediction of this trading relations view is that individuals who have greater asymmetries for faces (e.g., relative to baseline, greater performance advantage for faces presented to left than to right visual field, or greater activation in the right hemisphere in imaging) should show the converse for words on an individual-by-individual basis, depending on how the cooperation and competition play out during the course of development and experience. We also anticipate that some individuals will have more bilaterally graded representation and that there will be a large range of individual differences across the population. This

prediction is eminently testable through half-field studies, as well as functional imaging investigation, and we are currently undertaking such explorations including examining hemispheric asymmetries for faces and words within individual and across groups of young children, adolescents, and young adults (Dundas, Plaut, & Behrmann, 2011).

A further rather obvious prediction is that individuals with damage to the left VWFA and presenting with pure alexia might also be impaired at face recognition, relative to normal participants, albeit to a lesser extent than individuals with prosopagnosia following a lesion to the right FFA. The converse is also predicted: Individuals with a lesion to the right FFA and presenting with prosopagnosia might also be impaired at word recognition, relative to controls, albeit to a lesser extent than individuals with pure alexia following a lesion to the left VWFA. We have examined these predictions in a small group of individuals, all of whom were pre-morbidly normal and have acquired unilateral ventral cortex lesions (Behrmann & Plaut, 2011). In this study, we used the same series of face and word experiments to evaluate the performance of three adults with circumscribed unilateral right-hemisphere lesions and prosopagnosia and four pure alexic adults with circumscribed unilateral left-hemisphere lesions. Control participants matched to the two groups were also tested. In addition to the expected impairment in face recognition, the prosopagnosic individuals showed abnormal word recognition relative to the controls, albeit not as marked as in the pure alexics, and, in complementary fashion, the pure alexic individuals showed abnormal face recognition relative to the controls, albeit to a lesser extent than in the prosopagnosics (for related findings, see Buxbaum, Glosser, & Coslett, 1996, 1999). These empirical findings anchor a key prediction of our account, which is that hemispheric asymmetries for face and word recognition are graded and not fully and independently segregated.

### Limitations and extensions

The computational simulation presented here was intentionally kept as simple as possible in order to

provide the clearest illustration of the consequences of our putative computational principles for graded specialization of the neural representations of faces and words (and, to a more limited extent, houses). The most obvious limitations of the simulation are that it used a small set of highly schematized stimuli, a strictly feed-forward network architecture without lateral and top-down interactions within or between hemispheres, a lack of separate excitatory and inhibitory unit populations, and a biologically implausible learning procedure. Although we claim that the core findings regarding learned functional specialization do not depend critically on any of these simplifications, it is important to validate these findings in more realistic follow-up simulations. Such follow-up versions should use realistic stimuli, employ a more biologically plausible learning procedure (e.g., contrastive Hebbian learning; see O'Reilly, 1996), and permit only excitatory connections between layers by using a separate population of local inhibitory units within each layer. We do not anticipate that these elaborations will alter the basic operation of the model but they will bring it into much closer alignment with the operation of real neural systems.

### CONCLUSION

Our central hypothesis is that the commonalities in the neural mechanisms of face and word processing are not merely coincidental, as modular theories are left to conclude, but, rather, are the signature consequences of a set of general principles and constraints on neural computation that operate throughout cortex. We note that these principles themselves are not intended to be novel; in fact, we take them to be largely noncontroversial. Instead, the novelty derives from their common consequences for cortical organization and behaviour in two seemingly unrelated domains, specifically, in the context of words and faces,

Critically, when instantiated in explicit computational terms, these principles provide insight into why each of these properties is partial rather than absolute. This is because the principles and

constraints are inherently graded—adherence to the forces of cooperation, competition, and proximity in the process of learning cortical representations is a matter of degree as these constraints trade off against each other, and thus the consequences for neural and behavioural specialization are also graded. As a result, the implications of graded constraints go beyond explaining why neither pure alexia nor prosopagnosia is entirely “pure” and why, across a host of imaging studies, the FFA and VWFA show substantial responses to stimuli other than faces and words, respectively. They also imply that the functional and anatomical division between face and word recognition should be graded—despite the clear differences between the two domains, the FFA should be partially involved in word recognition, and the VWFA should be partially involved in face recognition. In this way, our theoretical perspective leads to important and otherwise unexpected predictions concerning the partial comingling of face and word processing, including face recognition impairments in pure alexia, word recognition impairments in prosopagnosia, graded participation of the FFA and VWFA in face/word recognition in normal observers, and a number of other implications that remain to be tested.

## REFERENCES

- Allison, T., McCarthy, G., Nobre, A. C., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers and colors. *Cerebral Cortex*, *5*, 544–554.
- Avidan, G., & Behrmann, M. (2009). Functional MRI reveals compromised neural integrity of the face processing network in congenital prosopagnosia. *Current Biology*, *19*, 1146–1150.
- Bach, F. R., & Jordan, M. I. (2002). Kernel independent component analysis. *Journal of Machine Learning Research*, *3*, 1–48.
- Baker, C. I., Liu, J., Wald, L. L., Kwong, K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences, USA*, *104*, 9087–9092.
- Bartlett, M. S., Movellan, J. R., & Sejnowski, T. J. (2002). Face recognition by independent component analysis. *IEEE Transactions on Neural Networks*, *13*, 1450–1464.
- Barton, J. J. S. (2008). Structure and function in acquired prosopagnosia: Lessons from a series of 10 patients with brain damage. *Journal of Neuropsychology*, *2*, 197–225.
- Barton, J. J. S. (2009). What is meant by impaired configural processing in acquired prosopagnosia? *Perception*, *38*, 242–260.
- Barton, J. J., Cherkasova, M. V., Press, D. Z., Intriligator, J. M., & O'Connor, M. (2004). Perceptual functions in prosopagnosia. *Perception*, *33*, 939–956.
- Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: Face-blind from birth. *Trends in Cognitive Science*, *9*, 180–187.
- Behrmann, M., & Kimchi, R. (2003). What does visual agnosia tell us about perceptual organization and its relationship to object perception? *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 19–42.
- Behrmann, M., Nelson, J., & Sekuler, E. (1998). Visual complexity in letter-by-letter reading: “Pure” alexia is not so pure. *Neuropsychologia*, *36*, 1115–1132.
- Behrmann, M., & Plaut, D. C. (2011). *Faces and words: Flip sides of the same brain*, Manuscript submitted for publication.
- Bilalic, M., Langner, R., Ulrich, R., & Grodd, W. (2011). Many faces of expertise: Fusiform face area in chess experts and novices. *Journal of Neuroscience*, *31*, 10206–10214.
- Blanz, V., & Vetter, T. (2003). Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*, 1063–1074.
- Bouvier, S. E., & Engel, S. A. (2006). Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex*, *16*, 183–191.
- Brem, S., Bach, S., Kucian, K., Guttorm, T. K., Martin, E., Lytinen, H., Brandeis, D., & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter-sound correspondences. *Proceedings of the National Academy of Science USA*, *107*, 7939–7944.
- Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., et al. (2006). Evidence for developmental changes in the visual word processing network beyond adolescence. *NeuroImage*, *29*, 822–837.

- Büchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature*, *394*, 274–277.
- Bukach, C., Le Grand, R., Kaiser, M. D., Bub, D. N., & Tanaka, J. W. (2008). Preservation of mouth region processing in two cases of prosopagnosia. *Journal of Neuropsychology*, *2*, 227–244.
- Burton, M. A., Bruce, V., & Hancock, P. J. B. (1999). From pixels to people: A model of familiar face recognition. *Cognitive Science*, *23*, 1–31.
- Burton, M. A., Young, A. W., Bruce, V., Johnston, R. A., & Ellis, A. W. (1991). Understanding covert recognition. *Cognition*, *39*, 129–166.
- Busigny, T., & Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*, *46*, 965–981.
- Buxbaum, L., Glosser, G., & Coslett, H. B. (1996). Relative sparing of object recognition in alexia–prosopagnosia. *Brain and Cognition*, *32*, 202–205.
- Buxbaum, L. J., Glosser, G., & Coslett, H. B. (1999). Impaired face and word recognition without object agnosia. *Neuropsychologia*, *37*, 41–50.
- Cai, Q., Lavidor, M., Brysbaert, M., Paulignan, Y., & Nazir, T. A. (2008). Cerebral lateralization of frontal lobe language processes and lateralization of the posterior visual word processing system. *Journal of Cognitive Neuroscience*, *20*, 672–681.
- Cantlon, J. F., Pineda, P., Dehaene, S., & Pelphrey, K. A. (2011). Cortical representations of symbols, objects, and faces are pruned back during early childhood. *Cerebral Cortex*, *21*, 191–199.
- Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. *NeuroImage*, *22*(1), 466–476.
- Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., 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.
- Cohen, L., Henry, C., Dehaene, S., Martinaud, O., Lehericy, S., Lemer, C., et al. (2004). The pathophysiology of letter-by-letter reading. *Neuropsychologia*, *42*, 1768–1780.
- Cohen, L., Lehericy, S., Henry, C., Bourgeois, M., Larroque, C., Sainte-Rose, C., et al. (2004). Learning to read without a left occipital lobe: Right-hemispheric shift of visual word form area. *Annals of Neurology*, *56*, 890–894.
- Cohen, L., Martinaud, O., Lemer, C., Lehericy, S., Samson, Y., Obadia, M., et al. (2003). Visual word recognition in the left and right hemispheres: Anatomical and functional correlates of peripheral alexias. *Cerebral Cortex*, *13*, 1313–1333.
- Cohen Kadosh, K., Cohen Kadosh, R., Dick, F., & Johnson, M. H. (2010). Developmental changes in effective connectivity in the emerging core face network. *Cerebral Cortex*, *21*, 1389–1394.
- Cohen Kadosh, K., & Johnson, M. H. (2007). Developing a cortex specialized for face perception. *Trends in Cognitive Science*, *11*, 367–369.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.
- Coslett, H. B., & Monsul, N. (1994). Reading with the right hemisphere: Evidence from transcranial magnetic stimulation. *Brain and Language*, *46*, 198–211.
- Cowey, A. (1979). Cortical maps and visual perception. The Grindley Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *31*, 1–17.
- Dailey, M. N., & Cottrell, G. W. (1999). Organization of face and object recognition in modular neural networks. *Neural Networks*, *12*, 1053–1074.
- Dailey, M. N., Cottrell, G. W., Padgett, C., & Adolphs, R. (2002). EMPATH: A neural network that categorizes facial expressions. *Journal of Cognitive Neuroscience*, *14*, 1158–1173.
- Damasio, A., Damasio, H., & Tranel, D. (1986). Prosopagnosia: Anatomical and physiological aspects. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young (Eds.), *Aspects of face processing* (pp. 279–290). Dordrecht, The Netherlands: Martinus Nijhoff.
- Davis, C. J. (1999). *The self-organizing lexical acquisition and recognition (SOLAR) model of visual word recognition*, Unpublished doctoral dissertation, University of New South Wales, Sydney, Australia.
- Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, *56*, 384–398.
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Science*, *15*, 254–262.
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Science*, *9*, 335–341.
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., et al. (2010). How learning to read changes the cortical networks for vision and language. *Science*, *330*, 1359–1364.

- de Schonen, S., Mancini, J., Camps, R., Maes, E., & Laurent, A. (2005). Early brain lesions and face-processing development. *Developmental Psychobiology*, *46*, 184–208.
- Devlin, J. T., Jamison, H. L., Gonnerman, L. M., & Matthews, P. M. (2006). The role of the posterior fusiform gyrus in reading. *Journal of Cognitive Neuroscience*, *18*, 911–922.
- De Yoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., et al. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National Academy of Sciences*, *93*, 2382–2386.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*, 2470–2473.
- Dundas, E., Plaut, D. C., & Behrmann, M. (2011). *The joint development of hemispheric lateralization for words and faces*. Submitted for publication.
- Epstein, R. A., Deyoe, E. A., Press, D. Z., Rosen, A. C., & Kanwisher, N. (2001). Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cognitive Neuropsychology*, *18*, 481–508.
- Epstein, R. A., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: Recognition, navigation, or encoding? *Neuron*, *23*, 115–125.
- Epstein, R. A., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601.
- Etemad, K., & Chellappa, R. (1997). Discriminant analysis for recognition of human face images. *Journal of the Optical Society of America A*, *14*, 1724–1733.
- Farah, M. J. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual object recognition. *Cognitive Neuropsychology*, *8*, 1–19.
- Farah, M. J. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of domain specificity in visual object recognition. *Current Directions in Psychological Science*, *1*, 164–169.
- Farah, M. J. (1996). Is face recognition “special”? Evidence from neuropsychology. *Behavioural Brain Research*, *76*, 181–189.
- Farah, M. J. (1999). Relations among the agnosias. In G. W. Humphreys (Ed.), *Case studies in the neuropsychology of vision* (pp. 181–200). Hove, UK: Psychology Press.
- Farah, M. J., O’Reilly, R. C., & Vecera, S. P. (1993). Dissociated overt and covert recognition as an emergent property of a lesioned neural network. *Psychological Review*, *100*, 571–588.
- Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 628–634.
- Feinberg, T. E., Schindler, R. J., Ochoa, E., Kwan, P. C., & Farah, M. J. (1994). Associative visual agnosia and alexia without prosopagnosia. *Cortex*, *30*, 395–412.
- Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (1999). Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron*, *24*, 205–218.
- Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2010). Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience*, *23*, 1723–1740.
- Gainotti, G., & Marra, C. (2011). Differential contribution of right and left temporo-occipital and anterior temporal lesions to face recognition disorders. *Frontiers in Human Neuroscience*, *5*, 1–11.
- Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition*, *103*, 322–330.
- Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. *Current Opinion in Neurobiology*, *11*, 219–224.
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*, 568–573.
- Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform “face area” is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*, 495–504.
- Golarai, G., Ghahremani, D. G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J. L., Gabrieli, J. D., et al. (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nature Neuroscience*, *10*, 512–522.
- Golarai, G., Liberman, A., Yoon, J. M., & Grill-Spector, K. (2010). Differential development of the

- ventral visual cortex extends through adolescence. *Frontiers in Human Neuroscience*, 3, 80.
- Grainger, J., Granier, J. P., Farioli, F., Van Assche, E., & van Heuven, W. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 865–884.
- Greenberg, A., Verstynen, T., Chiu, Y.-V., Yantis, S., Schneider, W., & Behrmann, M. (2011). *Spatiotopic structural connectivity underlying visual attention*, Manuscript submitted for publication.
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, 27, 649–677.
- Grill-Spector, K., Sayres, R., & Ress, D. (2006). High-resolution imaging reveals highly selective non-face clusters in the fusiform face area. *Nature Neuroscience*, 9, 1177–1185.
- Hanson, S. J., & Schmidt, A. (2011). High-resolution imaging of the fusiform face area (FFA) using multivariate non-linear classifiers shows diagnosticity for non-face categories. *NeuroImage*, 54, 1715–1734.
- Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, 106, 491–528.
- Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, 111, 662–720.
- Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Center-biased representation for characters in the human ventral visual stream. *Neuron*, 34, 479–490.
- Haxby, J. V. (2006). Fine structure in representations of faces and objects. *Nature Neuroscience*, 9, 1084–1086.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425–2429.
- Haxby, J. V., Petit, L., Ungerleider, L. G., & Courtney, S. M. (2000). Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *NeuroImage*, 11, 380–391.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, 3, 80–84.
- Hsiao, J., Shieh, D., & Cottrell, G. W. (2008). Convergence of the visual field split: Hemispheric modeling of face and object recognition. *Journal of Cognitive Neuroscience*, 20, 2298–2307.
- Ishai, A. (2008). Let's face it: It's a cortical network. *NeuroImage*, 40, 415–419.
- Ishai, A., Schmidt, C. F., & Boesiger, P. (2005). Face perception is mediated by a distributed cortical network. *Brain Research Bulletin*, 67, 87–93.
- Jacobs, R. A. (1997). Nature, nurture and the development of functional specializations: A computational approach. *Psychonomic Bulletin and Review*, 4, 299–309.
- Jacobs, R. A., & Jordan, M. I. (1992). Computational consequences of a bias toward short connections. *Journal of Cognitive Neuroscience*, 4, 323–336.
- Joseph, J. E., Gathers, A. D., & Bhatt, R. S. (2011). Progressive and regressive developmental changes in neural substrates for face processing: Testing specific predictions of the interactive specialization account. *Developmental Science*, 14, 227–241.
- Kanan, C. M., & Cottrell, G. W. (2010). Robust classification of objects, faces, and flowers using natural image statistics. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR-2010)*. Washington DC: IEEE Computer Society.
- Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, 3, 759–763.
- Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences, USA*, 107, 11163–11170.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17, 4302–4311.
- Kanwisher, N., Woods, R. P., Iacoboni, M., & Mazziotta, J. C. (1997). A locus in human extrastriate cortex for visual shape analysis. *Journal of Cognitive Neuroscience*, 9, 133–142.
- Kleinschmidt, A., & Cohen, L. (2006). The neural bases of prosopagnosia and pure alexia: Recent insights from functional neuroimaging. *Current Opinion in Neurology*, 19, 386–391.
- Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., & Gabrieli, J. D. E. (1989). Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 723–735.
- Kriegeskorte, N., Formisano, E., Sorge, B., & Goebel, R. (2007). Individual faces elicit distinct response

- patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences, USA*, *104*, 20600–20605.
- Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. *NeuroImage*, *21*, 946–953.
- Lavidor, M., Ellis, A. W., Shillcock, R., & Bland, T. (2001). Evaluating a split processing model of visual word recognition: Effects of word length. *Cognitive Brain Research*, *12*, 265–272.
- Levy, I., Hasson, U., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience*, *4*, 533–539.
- Levy, I., Hasson, U., & Malach, R. (2004). One picture is worth at least a million neurons. *Current Biology*, *14*, 996–1001.
- Liu, Y. C., Wang, A. G., & Yen, M. Y. (2011). “Seeing but not identifying”: Pure alexia coincident with prosopagnosia in occipital arteriovenous malformation. *Graefes’ Archive of Clinical and Experimental Ophthalmology*, *248*, 1087–1089.
- Mahon, B. Z., & Caramazza, A. (2011). What drives the organization of knowledge in the brain. *Trends in Cognitive Sciences*, *15*, 97–103.
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Science*, *7*, 293–299.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.
- McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159–188.
- McKone, E., & Robbins, R. (2011). Are faces special? In A. J. Calder, G. Rhodes, J. V. Haxby, & M. H. Johnson (Eds.), *Oxford handbook of face perception* (pp. 149–176). Oxford, UK: Oxford University Press.
- Mechelli, A., Gorno-Tempini, M. L., & Price, C. J. (2003). Neuroimaging studies of word and pseudo-word reading: Consistencies, inconsistencies, and limitations. *Journal of Cognitive Neuroscience*, *15*, 260–271.
- Moghaddam, B., Jebara, T., & Pentland, A. (2000). Bayesian face recognition. *Pattern Recognition*, *33*, 1771–1782.
- Montant, M., & Behrmann, M. (2000). Pure alexia: Review of cases. *Neurocase*, *6*, 265–294.
- Mozer, M. C. (1991). *The perception of multiple objects: A connectionist approach*. Cambridge, MA: MIT Press.
- Mozer, M. C., & Behrmann, M. (1990). On the interaction of selective attention and lexical knowledge: A connectionist account of neglect dyslexia. *Journal of Cognitive Neuroscience*, *2*, 96–123.
- Murtha, S., Chertkow, H., Beaugregard, M., & Evans, A. (1999). The neural substrate of picture naming. *Journal of Cognitive Neuroscience*, *11*, 399–423.
- Nelson, M. E., & Bower, J. M. (1990). Brain maps and parallel computers. *Trends in Neurosciences*, *13*, 403–408.
- Nestor, A., Behrmann, M., & Plaut, D. C. (2011). *The neural basis of visual word form processing—a multivariate investigation*, Manuscript submitted for publication.
- Nestor, A., Plaut, D. C., & Behrmann, M. (2011). Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proceedings of the National Academy of Science USA*, *108*, 9998–10003.
- Noppeney, U., Friston, K. J., & Price, C. J. (2003). Effects of visual deprivation on the organization of the semantic system. *Brain*, *126*, 1620–1627.
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Science*, *10*, 424–430.
- O’Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, *8*, 895–938.
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, *93*, 603–608.
- Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, *114*, 273–315.
- Petersen, S. E., & Fiez, J. (1993). The processing of single words studies with positron emission tomography. *Annual Review of Neurosciences*, *16*, 509–530.
- Petersen, S., Fox, P. T., Snyder, A., & Raichle, M. E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science*, *249*, 1041–1044.
- Plaut, D. C. (1999). A connectionist approach to word reading and acquired dyslexia: Extension to



- sequential processing. *Cognitive Science*, *23*, 543–568.
- Plaut, D. C. (2002). Graded modality-specific specialization in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology*, *19*, 603–639.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.
- Polk, T. A., & Farah, M. J. (2002). Functional MRI evidence for an abstract, not perceptual, word-form area. *Journal of Experimental Psychology: General*, *131*, 65–72.
- Polk, T. A., Stallcup, M., Aguirre, G. K., Alsop, D. C., D'Esposito, M., Detre, J. A., et al. (2002). Neural specialization for letter recognition. *Journal of Cognitive Neuroscience*, *14*, 145–159.
- Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *NeuroImage*, *19*, 473–481.
- Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Science*, *15*, 246–253.
- Price, C. J., & Mechelli, A. (2005). Reading and reading disturbance. *Current Opinion in Neurobiology*, *15*, 231–238.
- Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letter strings, and textures: A functional magnetic resonance imaging study. *The Journal of Neuroscience*, *16*, 5205–5215.
- Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology*, *74*, 1192–1199.
- Quartz, S. R., & Sejnowski, T. J. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences*, *20*, 537–596.
- Rajimehr, R., Young, J. C., & Tootell, R. B. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences, USA*, *106*, 1995–2000.
- Reddy, L., & Kanwisher, N. (2006). Coding of visual objects in the ventral stream. *Current Opinion in Neurobiology*, *16*, 408–414.
- Reich, L., Szwed, M., Cohen, L., & Amedi, A. (2011). A ventral visual stream reading center independent of visual experience. *Current Biology*, *21*, 363–368.
- Robertson, L. C., & Ivry, R. (2000). Hemispheric asymmetries: Attention to visual and auditory primitives. *Current Directions in Psychological Science*, *9*, 59–64.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533–536.
- Salvan, C. V., Ulmer, J. L., DeYoe, E. A., Wascher, T., Mathews, V. P., Lewis, J. W., et al. (2004). Visual object agnosia and pure word alexia: Correlation of functional magnetic resonance imaging and lesion localization. *Journal of Computer Assisted Tomography*, *28*, 63–67.
- Scherf, K. S., Behrmann, M., Humphreys, K., & Luna, B. (2007). Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Developmental Science*, *10*, 15–30.
- Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *Journal of Neuroscience*, *25*, 11055–11059.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889–893.
- Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. *Brain*, *115*, 15–36.
- Sewards, T. V. (2011). Neural structures and mechanisms involved in scene recognition: A review and interpretation. *Neuropsychologia*, *49*(3), 277–298.
- Shillcock, R., Ellison, M. T., & Monaghan, P. (2000). Eye-fixation behavior, lexical storage and visual word recognition in a split processing model. *Psychological Review*, *107*, 824–851.
- Spiridon, M., Fischl, B., & Kanwisher, N. (2006). Location and spatial profile of category-specific regions in human extrastriate cortex. *Human Brain Mapping*, *27*(1), 77–89.
- Spiridon, M., & Kanwisher, N. (2002). How distributed is visual category information in human occipitotemporal cortex? An fMRI study. *Neuron*, *35*(6), 1157–1165.
- Starrfelt, R., & Behrmann, M. (2011). Number reading in pure alexia: A review. *Neuropsychologia*, *49*(9), 2283–2298.
- Stephan, B. C., & Caine, D. (2009). Aberrant pattern of scanning in prosopagnosia reflects impaired face processing. *Brain and Cognition*, *69*, 262–268.

- Tagamets, M. A., Novick, J. M., Chalmers, M. L., & Friedman, R. B. (2000). A parametric approach to orthographic processing in the brain: An fMRI study. *Journal of Cognitive Neuroscience*, *12*, 281–297.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46A*, 225–245.
- Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2010). fMRI-adaptation studies of viewpoint tuning in the extrastriate and fusiform body areas. *Journal of Neurophysiology*, *103*, 1467–1477.
- Tong, M. H., Joyce, C. A., & Cottrell, G. W. (2008). Why is the fusiform face area recruited for novel categories of expertise? A neurocomputational investigation. *Brain Research*, *1202*, 14–24.
- Tootell, R. B. H., Mendola, J., Hadjikhani, H., Ledden, P., Liu, A., Reppas, J., et al. (1997). Functional analysis of V3a and related areas in human visual cortex. *Journal of Neuroscience*, *17*, 7060–7078.
- Tranel, D., Damasio, H., & Damasio, A. R. (1997). A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia*, *35*, 1319–1328.
- Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*, 71–86.
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *NeuroImage*, *16*, 765–780.
- Warrington, E. K., & Shallice, T. (1980). Word-form dyslexia. *Brain*, *103*, 99–112.
- Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology*, *50A*, 439–456.
- Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin and Review*, *8*, 221–243.
- Willems, R. M., Peelen, M. V., & Hagoort, P. (2010). Cerebral lateralization of face-selective and body-selective visual areas depends on handedness. *Cerebral Cortex*, *20*, 1719–1725.
- Wong, A. C., Gauthier, I., Woroch, B., DeBuse, C., & Curran, T. (2005). An early electrophysiological response associated with expertise in letter perception. *Cognitive, Affective and Behavioral Neuroscience*, *5*, 306–318.
- Yang, M.-H. (2002). Face recognition using kernel methods. In T. Diederich, S. Becker, & Z. Ghahramani (Eds.), *Advances in neural information processing systems* (Vol. 14, pp. 1457–1464). Cambridge, MA: MIT Press.
- Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology*, *15*, 2256–2262.
- Yovel, G., Tambini, A., & Brandman, T. (2008). The asymmetry of the fusiform face area is a stable individual characteristic that underlies the left-visual-field superiority for faces. *Neuropsychologia*, *46*, 3061–3068.