

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

David C. Plaut and Marlene Behrmann

Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213-3890

Address for correspondence:

David C. Plaut, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213-3890; email: plaut@cmu.edu; phone (412) 268-5145; fax (412) 268-2798

Running title: Complementary systems for faces and words

Acknowledgements: This work was funded by 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.

Abstract

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 favor the modular view, and yet close examination of those data reveal 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.

INTRODUCTION

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 behavioral processes, such as face, word, and object recognition, in a domain-specific manner (for recent reviews, see Kanwisher, 2010; McKone & Robbins, in press).¹ Consistent 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

1 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.

1
2
3 (e.g., Downing, Jiang, Shuman, & Kanwisher, 2001; Peelen & Downing, 2005; Schwarzlose,
4 Baker, & Kanwisher, 2005; Taylor, Wiggett, Downing, 2010; Willems, Peelen, & Hagoort, 2010)
5 and the visual word form area to orthographic inputs (e.g., Dehaene & Cohen, 2011). Indeed, in
6 each of these regions, the cortical response for the preferred category is about twice that for the
7 non-preferred category and this category-selectivity can be consistently observed in most
8 normal individuals, even across a range of very different experimental paradigms. All of this
9 attests to the robustness of the evidence that these regions are specialized for, and perhaps
10 even dedicated to, the recognition of particular object classes (Kanwisher, 2010; McKone &
11 Robbins, 2011).

12
13
14
15
16
17
18
19
20
21
22
23 The second approach recognizes the apparent selectivity of neural systems for certain
24 visual classes but argues that this selectivity need not implicate very specialized or dedicated
25 modules per se. This theoretical account entails one or both of two possible brain-behavior
26 organizations: rather than a single region alone subserving processing of a particular input type
27 (e.g., faces), multiple regions mediate the recognition of a particular object type and/or an
28 individual region mediates the neural representations of multiple object types. The claim, then,
29 is that, under either of these scenarios, specialization is more graded and regions may be
30 optimized for, but not necessarily dedicated to, a particular cognitive function. Consistent with
31 this alternative perspective, in addition to the FFA, multiple other cortical regions evince face-
32 selectivity, including the occipital face area (OFA, Gauthier et al., 2000), the posterior superior
33 temporal sulcus (Hoffman & Haxby, 2000) and the anterior temporal lobe (Kriegeskorte et al.,
34 2007; Rajimehr et al., 2009) and, indeed, multiple regions have sufficient neural information to
35 discriminate between individual face exemplars (Nestor, Plaut, & Behrmann, 2011) (for more
36 extended review, see Avidan & Behrmann, 2009; Haxby et al., 2000; Ishai, 2008). Furthermore,
37 it is not simply that the distributed network is domain-specific as there are now many fMRI
38 studies showing that even highly selective single regions, such as the FFA, evince a BOLD
39 response to different object classes, albeit with lesser degrees of activation than, for example, to
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 faces (e.g., Grill-Spector, Sayres, & Ress, 2006; Hanson & Schmidt, 2011; Haxby et al., 2001;
4
5 Haxby, Petit, Ungerleider, & Courtney, 2000; Ishai, Schmidt, & Boesiger, 2005; Nestor et al.,
6
7 2011; Norman, Polyn, Detre, & Haxby, 2006) and the same is true for the vWFA (Nestor,
8
9 Behrmann, & Plaut, submitted; Price & Devlin, 2011).

10
11
12 In this paper, we compare and contrast the more modular and more distributed accounts
13
14 with specific reference to two visual classes—faces and words. We choose these two classes
15
16 not only because, intuitively, they appear to be diametrically opposed but also because they
17
18 differ obviously along many other dimensions. Words and faces share little in common in their
19
20 overt geometry and so their image statistics share minimal, if any, overlap. Additionally,
21
22 whereas face representations are acquired naturally over the course of experience, word
23
24 recognition typically requires explicit instruction. Also, whereas faces are probably the most
25
26 ecologically relevant visual stimuli, orthographies have only been around for a few thousand
27
28 years and so the evolutionary trajectories of these two visual classes differ greatly.
29
30

31
32 We start by reviewing the clear evidence for the separability of the underlying systems
33
34 for words and faces. Thereafter, we present a proposal in which we argue that common
35
36 principles may account for both the similarities and differences in the mechanisms underlying
37
38 words and faces. We support this proposal with a computational simulation in which a common
39
40 underlying mechanism, constrained by a putative set of computational principles, mediates both
41
42 face and word recognition and demonstrates the types of functional specialization observed
43
44 empirically. Although we address the correspondences between brain and behavior in these two
45
46 particular domains, the argument has applicability to other aspects of cognition and its neural
47
48 correlates, as well, provided that these other cognitive behaviors place the same computational
49
50 demands on the visual recognition system. We also return to this point in the final discussion.
51
52

53 54 55 **Evidence for separability of word and face processing systems**

56
57 On a modular account of brain-behavior organization, words and faces engage separate
58
59
60

1
2
3 psychological and neural mechanisms and are, essentially, unrelated and independent. Support
4
5 for this view is substantial and gleaned from functional imaging investigations, as well as from
6
7 neuropsychological studies (Kleinschmidt & Cohen, 2006)..
8
9

10 11 12 *The Visual Word Form Area*

13
14 Numerous functional imaging studies have demonstrated that the word module or 'visual
15
16 word form area' (VWFA) (e.g., Cohen et al., 2000; Cohen et al., 2003; Dehaene, Cohen,
17
18 Sigman, & Vinckier, 2005; Dehaene & Cohen, 2011) responds selectively to visually presented
19
20 words and letter strings (e.g., Fiez, Balota, Raichle, & Petersen, 1999; Mechelli, Gorno-Tempini,
21
22 & Price, 2003; Petersen, Fox, Snyder, & Raichle, 1990; Petersen & Fiez, 1993; Turkeltaub,
23
24 Eden, Jones, & Zeffiro, 2002) to a greater degree than digits (Polk et al., 2002) or pseudo-letters
25
26 (Allison, McCarthy, Nobre, Puce, & Belger, 1994; Cohen & Dehaene, 2004), but not to spoken
27
28 words (Cohen & Dehaene, 2004). The VWFA activation is located in left extrastriate cortex
29
30 (Talairach coordinates: x=-43, y=-54, z=-12), is identifiable in single subjects (Puce, Allison,
31
32 Asgari, Gore, & McCarthy, 1996) and is sensitive to the individual's experience – Hebrew
33
34 readers show greater activation of this region for Hebrew than English words and vice versa
35
36 (Baker et al., 2007), and activation in this area is correlated with literacy (Dehaene et al., 2007,
37
38 2010; Dehaene & Cohen, 2011). The VWFA is situated anterior to retinotopic cortex and,
39
40 consistent with this, activation is relatively insensitive to retinal position and to the font, size or
41
42 case of the input (Polk & Farah, 2002). Activation of VWFA, as measured in ERPs, is rapid,
43
44 emerging around 150-200ms after stimulus onset (McCandliss, Cohen, & Dehaene, 2003). In
45
46 normal readers, the minimal increase in reaction time (RT) as a function of word length (Lavidor,
47
48 Ellis, Shillcock, & Bland, 2001; Weekes, 1997) is attributed to the parallel processing of multiple
49
50 letters (to the limits of foveal acuity, i.e., around 9 letters) and this parallel processing is ascribed
51
52 to the functionality of the VWFA.
53
54
55
56

57 Further support for the circumscribed functionality of the VWFA comes from studies of
58
59
60

1
2
3 premorbidly literate individuals with 'pure alexia' (for review of cases, see Montant & Behrmann,
4 2000; Starrfelt & Behrmann, 2011). The lesion site in these cases is typically in the left
5
6 2000; Starrfelt & Behrmann, 2011). The lesion site in these cases is typically in the left
7 occipitotemporal area along the fusiform and adjacent lingual gyrus, with possible incursion to
8
9 the inferior longitudinal fasciculus (Cohen, Henry et al., 2004; Cohen et al., 2003; Feinberg,
10
11 Schindler, Ochoa, Kwan, & Farah, 1994; Salvan et al., 2004) and overlaps the region of the
12
13 VWFA activation reported above (Hasson, Levy, Behrmann, Hendler, & Malach, 2002; Petersen
14
15 et al., 1990; Puce et al., 1996). The characteristic profile of pure alexia is a linear increase in RT
16
17 as a function of the number of letters in the input (giving rise to the label 'letter-by-letter reading')
18
19 and this is assumed to reflect the breakdown of parallel processing in the VWFA and the
20
21 subsequent reliance on a serial, laborious left-right letter spelling strategy (McCandliss et al.,
22
23 2003; Warrington & Shallice, 1980). The patients are not aphasic, typically showing intact
24
25 production and comprehension of spoken language along with normal writing, all of which
26
27 supports the circumscribed nature of the problem as a specific difficulty in processing visual
28
29 word forms (but see Starrfelt & Behrmann, 2011, for discussion of high association with an
30
31 impairment of number processing as well).
32
33
34
35
36
37

38 *The Fusiform Face Area*

39
40 Just as in the case of the VWFA, there is substantial evidence for face-processing
41
42 specificity gleaned from fMRI studies and from patient studies. Functional imaging studies have
43
44 provided evidence that the region that is functionally specialized for faces, the 'fusiform face
45
46 area' (FFA) ($x=40, y=-55, z=-10$), is selectively activated by faces, especially upright faces,
47
48 over other non-face objects (Kanwisher, 2010; Puce et al., 1995; Yovel & Kanwisher, 2005) and
49
50 over animal or cartoon faces (e.g., Kanwisher, 2000; Kanwisher et al, 1997; Kanwisher, Woods,
51
52 Iacoboni, & Mazziotta, 1997; Sergent, Ohta, & MacDonald, 1992a, 1992b; Spiridon, Fischl, &
53
54 Kanwisher, 2006) and the magnitude of the activation is correlated with face identification ability
55
56 (Furl, Garrido, Dolan, Driver, & Duchaine, 2010; Yovel, Tambini, & Brandman, 2008). FFA
57
58
59
60

1
2
3 activation is situated anterior to retinotopic cortex and, consistent with this, activation is
4
5 relatively insensitive to retinal position and to size, color, format (drawing or photographs) and
6
7 viewpoint of input. The FFA is selectively activated for faces but abuts other cortical regions
8
9 that are specialized for other visual categories, such as scenes, animals and tools (e.g., Reddy
10
11 & Kanwisher, 2006; Spiridon & Kanwisher, 2002; also, Puce et al., 1995, 1996; Tranel,
12
13 Damasio, & Damasio, 1997).

14
15
16 Correspondingly, lesions to the FFA (Bouvier & Engel, 2006; Damasio, Damasio, &
17
18 Tranel, 1986; Kleinschmidt & Cohen, 2006) result in prosopagnosia, a selective impairment in
19
20 face recognition. The lesion in prosopagnosia is often bilateral, affecting the temporo-occipital
21
22 cortex in the region of the FFA, but unilateral right hemisphere lesions to this same region may
23
24 suffice to give rise to this disorder (Barton, 2008; Bouvier & Engel, 2006), and prosopagnosia
25
26 can also be congenital or developmental in the absence of a frank lesion (Behrmann & Avidan,
27
28 2005). The difficulty in recognizing faces can be dramatic, including failures to recognize friends
29
30 or even close family members. Unlike normal observers, these individuals do not obviously
31
32 exhibit the advantage for upright over inverted faces (occasionally even showing an inversion
33
34 superiority effect; Farah, 1996; Farah, Tanaka, & Drain, 1995) and do not appear to process
35
36 faces configurally, thus failing to evince the benefit from the presence of the whole face over just
37
38 parts of the face (Barton, 2009; Barton, Cherkasova, Press, Intriligator, & O'Connor, 2004;
39
40 Busigny & Rossion, 2010; Tanaka & Farah, 1993).

41
42
43 Taken together, these studies provide empirical support for the claim that there is
44
45 specialized processing of faces and words associated with two distinct cortical modules, the
46
47 FFA for faces and the VWFA, for words, and that these two systems are separable and
48
49 independent.
50
51

52 53 54 55 **Not only differences but also commonalities**

56
57 Although there is general consensus that the FFA and VWFA are tuned to faces and words,
58
59
60

1
2
3 respectively, there are also intriguing empirical data that suggest that both their tuning and their
4 hemispheric specialization is relative or graded. For example, it appears that both the VWFA
5 and FFA can be activated by a wide range of stimuli, not just faces or words: the VWFA is
6 strongly activated in response to checkerboards, pictured objects, and verb naming to pictures
7 (Devlin et al., 2006; Murtha, Chertkow, Beauregard, & Evans, 1999; Price & Devlin, 2003, 2011)
8 and even to non-visual inputs such as Braille (Büchel, Price, Friston, 1998; Reich, Szwed,
9 Cohen, & Amedi, 2011), whereas the FFA is activated by a range of non-face stimuli, such as
10 houses and cars, but also novel objects such as Greebles (Gauthier et al., 1999) and chess
11 configurations (Bilalic et al., 2011), although the full extent of the selectivity is still controversial
12 (Gauthier et al., 1999; Grill-Spector et al., 2006; Haxby, 2006). Neither the FFA nor the VWFA,
13 however, appears to be as strongly activated by these other stimuli as is the case when shown
14 the 'preferred' input type, reflecting perhaps the graded nature of the underlying
15 representations. Recent imaging studies adopting multivariate methods applied to fMRI data of
16 ventral visual cortex have begun to uncover the co-mingling of patterns of activation associated
17 with different stimulus types (e.g., face and word representations) to an even greater degree
18 than was revealed in earlier studies employing univariate analyses (for an example of a recent
19 study using multi-voxel pattern analysis, see Nestor et al., 2011). In these multivariate studies, it
20 is not simply the magnitude of the activation that is crucial but the distribution of the neural
21 information in the patterns of voxel activation.

22
23 Also relevant to the similarities across classes is the observation that almost all fMRI
24 and ERP studies show bilateral activation for words and for faces, albeit with differential
25 hemispheric asymmetry and greater scalp potential for the preferred stimulus type in the
26 corresponding hemisphere – words on the left and faces on the right (e.g., see Hasson et al.,
27 2002; Kanwisher et al., 1997; Kronbichler et al., 2004; Price & Mechelli, 2005; Puce et al., 1996;
28 Sergent et al., 1992; Tagamets, Novick, Chalmers, & Friedman, 2000). Moreover, the peak
29 activation for words in vWFA and for faces in FFA (although coordinates differ a little across
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 different studies) are very comparable in the two hemispheres (for example, Talairach
4 coordinates for peak for words $x=-43$, $y=-54$, $z=-12$ and for faces $x=40$, $y=-55$, $z=-10$). We also
5 note that these coordinates roughly demarcate cortical sites that are anterior to retinotopic
6 cortex but are situated in what would be the anterior extrapolation of the fovea (Levy et al.,
7 2001; Hasson et al., 2002). The localization of the functional regions in this cortical location is
8 consistent with the invariance of face and word activation over retinal position of the inputs, but
9 also with the fact that reliance on fine-grained visual discrimination is a necessary component of
10 both face and word recognition.
11
12
13
14
15
16
17
18
19

20 Somewhat surprisingly, there has not been a systematic examination of the word
21 recognition of prosopagnosic individuals and the face recognition of pure alexic individuals.
22 There are some hints, however, that each hemisphere may play a dual, albeit graded, role in
23 both face and word recognition. For example, it has been reported that the face recognition
24 impairment is more severe following bilateral than unilateral lesions (Damasio et al., 1985;
25 Gainotti & Marra, 2011), implicating both hemispheres to some extent and that TMS of the RH
26 even impairs reading in patients with LH lesions (Coslett & Monsul, 1994).
27
28
29
30
31
32
33
34
35

36 Additionally, in a few case studies in which both stimulus classes have been examined,
37 some prosopagnosic individuals show increased word length effects in reading aloud single
38 words: for example, the slope of the reaction time in single word reading was 104 ms and 241
39 ms per additional letter for prosopagnosic patients SM and RN, respectively, compared with the
40 normal slope of about 10 ms for words 3 through 8 letters in length (Behrmann & Kimchi, 2003).
41 In a complementary fashion, there have only been a few reports of pure alexic individuals who
42 have difficulties with face recognition (also see Farah 1991, 1992, 1999, for listing of co-
43 occurrences of different forms of agnosia/alexia) although this is not always assessed in these
44 cases. One recent relevant study documents a case with left occipital arteriovenous
45 malformation in whom both pure alexia and prosopagnosia were evident (Liu et al., 2011). Many
46 studies do report abnormalities in the recognition of non-orthographic stimuli in pure alexia even
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 after a unilateral lesion (Behrmann, Nelson & Sekuler, 1998; Starrfelt & Behrmann, 2011) and
4
5 so one might predict that a decrement in face recognition in these cases, as well. As evident,
6
7 closer scrutiny of the existing data, to the extent they are available, suggest that there may be
8
9 more overlap in face and word processing in the preeminent face (right FFA) and word (left
10
11 VWFA) regions than originally considered. To account for both the apparent differences and the
12
13 similarities, we propose an account that differs from the strictly modular or domain-specific view.
14
15
16
17

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

19
20 The theoretical proposal outlined in this paper adopts an alternative perspective with respect to
21
22 the key systems engaged in face and word processing. The central idea is that visual object
23
24 recognition (e.g., face and word recognition) is supported, not by highly specialized (or
25
26 dedicated) modules per se, but by a distributed and interactive network of brain regions with
27
28 similar computations but whose organization is strongly shaped and modified by experience.
29
30 This view then incorporates both the claim that multiple cortical regions are engaged and that
31
32 these nodes of this distributed network play a role in representing more than one stimulus type.
33
34 Importantly, on this view, the functional specialization of brain regions is graded rather than
35
36 absolute and reflects the consequences of a set of general principles and constraints on neural
37
38 computation that operate throughout cortex but that nonetheless have distinct implications for
39
40 different classes of stimuli. Note that, on this account, there is no appeal to pre-specified
41
42 modules and, rather than claiming de facto sensitivity to different visual classes, the origin and
43
44 emergence of these graded mechanisms is captured too. The novelty of this approach is not the
45
46 principles themselves (see both Dehaene & Cohen, 2011, and Price & Devlin, 2011, for similar
47
48 notions about the VWFA) but their integrated application to derive common consequences for
49
50 cortical organization and behavior for words and faces.
51
52
53
54

55 This alternative proposal takes as its starting assumptions three general principles of
56
57 neural computation, that the neural system for face/word recognition is distributed, that
58
59
60

1
2
3 knowledge is represented in this system by cooperation and competition between the
4
5 processing units, and that the organization of the system is constrained by topographical
6
7 considerations, pressure for proximity and the division of labor between the two hemispheres of
8
9 the brain. We expand on these assumptions here.
10

11 12 13 14 *Distributed representation and knowledge* 15

16 We assume that the neural system for visual object recognition consists of a set of
17
18 hierarchically organized cortical areas, ranging from local retinotopic information in V1 through
19
20 more global, object-based and semantic information in anterior temporal cortex (Grill-Spector &
21
22 Malach, 2004). At each level, the visual stimulus is represented by the activity of a large
23
24 number of neurons, and each neuron participates in coding a large number of stimuli.
25
26 Generally, stimuli that are similar with respect to the information coded by a particular region
27
28 evoke similar (overlapping) patterns of activity. The set of constraints on how activity at one
29
30 level produces activity at the next level—that is, the knowledge of how features combine to form
31
32 features at the next level—is encoded by the pattern of synaptic connections and strengths
33
34 between and within the regions. Learning involves modifying these synapses in a way that
35
36 alters the representations evoked by visual stimuli—typically in a way that captures the relevant
37
38 information in the domain better and that supports more effective behavioral outcomes. With
39
40 extended experience, expertise develops through the refinement, specialization, and elaboration
41
42 of representations, requiring the recruitment of additional neurons and a larger region of cortex
43
44 (Quartz & Sejnowski, 1997).
45
46
47
48
49
50

51 *Representational cooperation and competition* 52

53 As illustrated by artificial neural networks (e.g., McClelland & Rumelhart, 1985), a single
54
55 pattern of synaptic connections can learn to encode the knowledge needed to represent many
56
57 stimuli, but its ability to do so depends on the degree to which the relevant knowledge is
58
59
60

1
2
3 consistent or systematic (i.e., similar representations at one level correspond to similar
4 representations at another). In general, systematic domains benefit from highly overlapping
5 neural representations that support generalization, whereas unsystematic, unrelated domains
6 require largely non-overlapping representations to avoid interference. Thus, if a cortical region
7 represents one type of information, it is ill-suited to represent another type of information that
8 requires unrelated knowledge, and so, that information must be represented by a different
9 region. On the other hand, effective cognitive processing requires the coordination of multiple
10 levels of representation within a given domain, and often across multiple domains. Of course,
11 representations can cooperate directly only to the extent they are connected—that is, there are
12 synapses between the regions encoding the relevant knowledge of how they are related;
13 otherwise, they must cooperate indirectly through mediating representations. In this way, the
14 neural organization of cognitive processing is strongly constrained by available connectivity (see
15 Mahon & Caramazza, 2011, for a similar argument regarding connectivity serving as an
16 endogenous constraint on topographic organization in the ventral stream).
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 *Topography, proximity, and hemispheric organization*

37
38 Brain organization must permit sufficient connectivity among neurons to carry out the
39 necessary information processing, but the total axonal volume must fit within the confines of the
40 skull (for similar discussion, see Cowey, 1979). This constraint is severe: if the brain's 10^{11}
41 neurons were placed on a sphere and fully interconnected with 0.1 mm radius axons,
42 accommodating the axon volume would require a sphere over 20 km in diameter (Nelson &
43 Bower, 1990). If we think of brain organization as the result of a complex optimization process
44 that minimizes “costs” associated with the degree to which various pressures or biases are
45 violated, then clearly there is a strong pressure to keep connectivity as local as possible. Long-
46 distance projections are certainly present in the brain but they are relatively rare and
47 presumably play a sufficiently critical functional role to offset their cost in volume. In fact, the
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 organization of human neocortex as a folded sheet can be understood as a compromise
4
5 between the spherical shape that would minimize long-distance axon length and the need for
6
7 greater cortical area to support highly elaborated representations. The organization into two
8
9 hemispheres is also relevant here, as interhemispheric connectivity is largely restricted to
10
11 homologous areas and is thus vastly less dense than connectivity within each hemisphere.
12
13 Even at a local scale, the volume of connectivity within an area can be minimized by adopting a
14
15 topographic organization so that related information is represented in as close proximity as
16
17 possible (Jacobs & Jordan, 1992). This is seen most clearly in the retinotopic organization of
18
19 early visual areas, given that light falling on adjacent patches of the retina is highly likely to
20
21 contain related information. Note that the dimensions of this topography are not in the
22
23 Cartesian (x, y) coordinates that apply naturally to images, but something closer to polar (r, θ)
24
25 coordinates, where eccentricity (central vs. peripheral) is coded along one axis and rotational
26
27 angle is coded along another (e.g., De Yoe et al., 1996; Grill-Spector & Malach, 2004; Sereno et
28
29 al., 1995; Tootell et al., 1997). The relevant dimensions of similarity for higher-level visual areas
30
31 are, of course, far less well understood, but the local connectivity constraint is no less pertinent
32
33 (Jacobs, 1997).
34
35
36

37
38 Despite these commonalities, the principles rule out using the very same cortical region
39
40 to represent both faces and words because these stimuli require entirely distinct primitives to be
41
42 represented as visual objects, and typically have distinct consequences for cognition (faces
43
44 designate individuals, whereas—apart from proper names—words designate objects, actions,
45
46 properties, typically at a basic rather than individual level). Given the need for written words to
47
48 interact with aspects of language that are left lateralized in most individuals (Cai et al., 2008), it
49
50 follows from representational competition and cooperation that visual word representations
51
52 would be predominantly located in the dominant language hemisphere (Price & Devlin, 2011)
53
54 whereas face representations would be located in the homologous region in the RH. Indeed,
55
56 some data to support this competition/cooperation comes from the observation that, with
57
58
59
60

1
2
3 increasing literacy, there is a decrease in response to faces in the VWFA (Dehaene et al., 2010)
4
5 and that, in four-year olds, performance in identifying alphanumeric characters (digits and
6
7 letters) was correlated with a decrease in left fusiform activity (Cantlon et al., 2011). In both of
8
9 these studies, however, the competition appears to be restricted to the left hemisphere with the
10
11 trading relations between faces and words manifesting in the left fusiform region. The
12
13 hypothesis we propose, however, encompasses the left and the right hemispheres with the
14
15 competition and cooperation playing out for words and faces across both sides of cortex.
16
17
18
19

20 21 **SIMULATION**

22
23 To date, the majority of computational work on face recognition has an applied focus with only
24
25 tangential relevance to the human cognitive and neural system—this includes approaches
26
27 based on Principal Components Analysis (e.g., Turk & Pentland, 1991), Independent
28
29 Components Analysis (e.g., Bartlett et al., 2002), Linear Discriminant Analysis (e.g., Etemad &
30
31 Chellapa, 1997), Kernel methods (e.g., Bach & Jordan, 2002; Yang, 2002), 3D morphable
32
33 models (e.g., Blanz & Vetter, 2003) and Bayesian inference (e.g., Moghaddam et al., 2000).
34
35 Modeling efforts that explicitly address psychological and neuropsychological issues (e.g.,
36
37 Burton et al., 1991; Farah et al., 1993) have tended to focus on the interaction of higher-level
38
39 knowledge with rather less consideration of low- and intermediate-level visual representation
40
41 and processing (although see Burton et al., 1999). More recently, Cottrell and colleagues
42
43 (Dailey & Cottrell, 1999; Dailey et al., 2002; Hsiao et al., 2008; Kanan & Cottrell, 2010; Tong et
44
45 al., 2008) have extended this work by coupling distributed network modeling with more realistic
46
47 assumptions about early visual processing.
48
49

50
51 A similar situation holds with regard to word recognition. Although some early cognitive
52
53 and neuropsychological modeling employed hierarchical visual representations of letters and
54
55 words (Mozer, 1991; Mozer & Behrmann, 1990; McClelland & Rumelhart, 1981), the vast
56
57 majority of more recent work has emphasized higher-level interactions of orthographic,
58
59
60

1
2
3 phonological, and semantic knowledge (e.g., Coltheart et al., 2001; Harm & Seidenberg, 1999,
4 2004; Plaut et al., 1996; Perry et al., 2007; although see Plaut, 1999). Efforts to model
5 orthographic representations per se (e.g., SOLAR, Davis, 1999; SERIOL, Whitney, 2001) have
6 typically focused more narrowly on letter position effects in orthographic priming. One notable
7 exception is the split-fovea model (Shillcock et al. 2000), which explicitly considers the
8 representational implications of a divided visual field. Although this specific model runs into
9 some empirical difficulties (see, e.g., Grainger et al., 2006), there is no doubt that the cortical
10 representation of words is shaped in important ways by hemispheric organization and
11 specialization (Cai et al., 2008).

12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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 modeling 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 these 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 faces 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

1
2
3 faces. Due to the graded nature of the learning constraints, this specialization should be only
4 partial, with both regions participating in processing both types of stimuli to a certain degree.
5
6 Houses are expected to be represented by more peripheral regions in fusiform cortex,
7
8 analogous to the “parahippocampal place area” (PPA; Epstein & Kanwisher, 1998; Levy et al.,
9
10 2001).
11
12

13
14 All simulations were developed within the Light Efficient Network Simulator (Lens;
15
16 version 2.63), developed by Doug Rohde and available for download at tedlab.mit.edu/~dr/Lens.
17
18
19

20 21 **Methods**

22
23 *Stimuli.* As the goals of the current work are to explore and illustrate the implications of
24 a set of putative computational principles rather than to build a realistic model of visual face and
25 word perception, the task and network architecture employed in the simulation were kept as
26 simple as possible. The stimuli used in the simulation were derived from 32x32 bit schematic
27 line drawings of faces, houses and words that embodied critical differences in the demands of
28 recognition of these classes of stimuli (see Figure 1). Each of 34 faces differed in terms of small
29 changes in the positions or shapes of central features (e.g., separation and height of eyes,
30 length of nose, height and width of mouth). Each of 40 3-letter (CVC) words were created from
31 combinations of 5 possible letters for each position and, like faces, differed from each other only
32 in terms of features within central vision. By contrast, each of 9 houses differed in terms of
33 properties that varied across the entire visual field (e.g., size of windows, number of eaves,
34 presence of porch, size of base). Note that these rather small differences place high demands
35 on fine visual acuity to ensure accurate discrimination between exemplars.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 Insert Figure 1 approximately here

52
53 Each item was presented at nine different scales, ranging from 1.0 to 0.6 in steps of 0.05, for a
54 total of 747 input patterns. For each pattern, retinotopic input activation was generated by
55 smoothing the original bit patterns by convolving them with a Gaussian (sd=0.5) and then
56
57
58
59
60

1
2
3 transforming the resulting values into polar coordinates (r, θ) , such that eccentricity r varied
4 along the horizontal axis (with central information on the left and peripheral information on the
5 right), and visual angle θ varied along the vertical axis (see Figure 2 for examples).
6
7
8

9
10 Insert Figure 2 approximately here

11
12 As the current work is concerned only with the nature of the visual representations of
13 various stimulus classes, no attempt was made to approximate the structure of higher-level
14 information that such representations provide access to, beyond the need to identify
15 (individuate) each unique face, word, and house (despite changes in scale). Accordingly, the
16 output representations used in the simulation consisted of individual "localist" units for each of
17 the 34 faces, 40 words, and 9 houses. We recognize, of course, that this is unlikely to be the
18 veridical output of the system.
19
20
21
22
23
24
25
26
27
28

29 *Network architecture.* The network architecture is depicted in Figure 3. In the model,
30 32x32 retinotopic visual input to each hemisphere (in polar coordinates) is mapped via 64
31 (32x2) intermediate units in each hemisphere (corresponding to fusiform cortex) onto a set of 83
32 "Identity" units (one for each unique word, face and house). In addition, to approximate the
33 influence of a left-hemisphere specialization for language, word inputs were also trained to
34 activate one of a set of 40 "Language" units that receive input solely from the left-hemisphere
35 intermediate units. Finally, the input-to-intermediate connections were subject to strong
36 topographic bias during learning. Although this bias is assumed to be enforced in the brain by
37 the relative density of synapses as a function of distance, the small scale of the simulation made
38 it more appropriate to implement this bias slightly differently. Specifically, the input units were
39 fully connected to the intermediate units, but the efficacy of learning decreased as a Gaussian
40 function ($sd=3.0$) of the distance between the connected units (Plaut, 2002). As we were
41 primarily concerned with the impact of eccentricity on learning, this metric considered only
42 horizontal distance in the simulation (i.e., all units in the same column in Figure 2 were
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 considered to have the equivalent functional position). Thus, in practical terms, learning was
4
5 effective on connections from inputs directly “below” a given intermediate unit, but increasingly
6
7 ineffective on connections from units with progressively different horizontal positions. For similar
8
9 reasons, although we would claim that all connectivity in the brain is subject to a topographic
10
11 bias, we did not apply this bias to any of the intermediate-to-output connections because we
12
13 had no hypothesis concerning the relative proximity of semantic or identity information (beyond
14
15 the left-lateralization of language information). The important consequence of this is that the
16
17 Identity units for faces, words and houses are equivalent in their connectivity with the
18
19 intermediate units, and thus any distinction in the specialization of the intermediate units must
20
21 arise solely from properties of the inputs. Finally, the simulation employed a feed-forward
22
23 architecture, without lateral (within-layer) or top-down connections, solely for computational
24
25 convenience.
26
27
28

29
30 Insert Figure 3 approximately here
31
32

33
34 *Training and Testing.* When presented with a scaled version of each face, word or
35
36 house, the network was trained to activate the correct Identity unit (and, for words, the correct
37
38 “Language” unit). Back-propagation (Rumelhart et al., 1986) was used to calculate how to
39
40 change each connection weight in the network to reduce the discrepancy between the output
41
42 activation pattern generated by the network and the correct pattern. Although not biologically
43
44 plausible in literal form, back-propagation is functionally equivalent to more plausible procedures
45
46 such as contrastive Hebbian learning (see, e.g., O'Reilly, 1996). The topographic bias on
47
48 learning at the intermediate layer was implemented by scaling these weight changes by a
49
50 decreasing (Gaussian) function of the horizontal distance between the connected input and
51
52 intermediate units. Following training, the network was considered to be correct if, for a given
53
54 input, the correct Identity unit was more active than any other.
55
56

57
58 The primary basis for establishing specialization in the network was its performance
59
60

1
2
3 following lesions restricted to spatially contiguous areas of either the left or right intermediate
4 (fusiform) layer. Lesions were administered by removing three adjacent columns (6 units in
5 total) from one of these layers and evaluating the performance of the damaged network for each
6 of the 747 input patterns (83 identities x 9 scales). The horizontal position of these lesions were
7 varied systematically in order to evaluate the relative specialization of each intermediate layer
8 for each stimulus class as a function of visual eccentricity.
9
10
11
12
13
14
15
16
17
18

19 **Results and Discussion**

20 After 1000 training presentations of each pattern, the network is fully accurate at recognizing
21 instances of each face, word and house (see Figure 4). Over the course of acquisition,
22 performance on houses is better than on the other stimulus classes because there are fewer of
23 them to differentiate. Performance on words is better than on faces in part because the latter
24 involve more subtle featural distinctions and, in part, because the extra demands of activating
25 “Language” information for words provides additional error (and therefore learning). By the end
26 of training, however, performance on all three classes is equivalent and at ceiling.
27
28
29
30
31
32
33
34
35

36 Insert Figure 4 approximately here
37
38
39

40 To illustrate the effects of the topographic bias on learning in the network, Figure 5
41 shows examples of the “receptive” and “projective” fields learned by two intermediate units. The
42 left display is for a unit that has a receptive field in central vision (i.e., toward the left of the
43 retinotopic input) and has output that is largely selective for faces (third and fourth rows in the
44 top group of units). By contrast, the right display is for a unit that has a more peripheral
45 receptive field (i.e., toward the center or right of the input) and is largely selective for houses
46 (last row in the top group). These weight diagrams illustrate the impact of the topographic
47 constraint on learning, and provide indirect evidence for learned category specificity of
48 intermediate units as a function of their eccentricity.
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Insert Figure 5 approximately here

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 et al., 2001) and our derived predictions for graded functional specialization.²

Insert Figure 6 approximately here

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, 2004). Competition between inconsistent information (and cooperation between word representations and language-related information) lead to substantial but still graded hemispheric specialization, with words represented primarily on the left and faces primarily on the right.

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

GENERAL DISCUSSION

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

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

1
2
3 represented over separate regions to avoid interference; and

- 4
5 3. *topography, proximity, and hemispheric organization*: representational
6
7 cooperation must be accomplished with largely local connectivity between
8
9 topographically organized brain regions, and with limited connectivity between
10
11 hemispheres, so that total axonal volume fits within the confines of the skull.
12
13

14 These constraints have a fundamental impact on how faces and words are represented
15
16 and processed within the visual system.
17

18 Here, we show that a small-scale simulation that is trained to recognize faces, words
19
20 and houses in a manner consistent with these principles illustrates how a topographic constraint
21
22 on learning can give rise to learned category specificity of intermediate units as a function of
23
24 their eccentricity. This topographic constraint is further anchored by the reliance on fine-grained
25
26 visual processing for discriminating subtle visual differences among words and among faces.
27
28 Specifically, left-hemisphere lesions to the region of the intermediate layer near central visual
29
30 information (analogous to the VWFA) produce a marked impairment in word recognition, but
31
32 also a milder impairment in face recognition (relative to house recognition). By contrast, lesions
33
34 to the corresponding region in the right hemisphere (analogous to the FFA) impair face
35
36 recognition most, but also word recognition to a lesser extent. These findings are in qualitative
37
38 agreement with existing observations and our derived predictions for graded functional
39
40 specialization.
41
42
43

44 The results of the simulation provide an existence proof of a system in which face and
45
46 word recognition are subject to the same computational constraints but in which relative
47
48 specialization, by virtue of other competing pressures (to restrict connection length specifically
49
50 with language areas, see Cai et al., 2008; Price & Devlin, 2011) also demonstrates some
51
52 functional specialization. The idea that there are both many similarities as well as differences
53
54 among the mechanisms supporting face and word recognition is endorsed by the existing
55
56 neuroimaging studies (many or even most of which reflect bilateral activation for faces and
57
58
59
60

1
2
3 words) and some neuropsychological studies, which show dual impairment following a unilateral
4
5 hemispheric lesion.
6
7

8 9 10 **Converging evidence**

11
12 There are some additional considerations, which favor a common-mechanism perspective. Both
13
14 face and word recognition are domains with which most individuals have extensive experience
15
16 and expertise. Both classes place demands on high-acuity information to encode subtle but
17
18 critical visual information and thus, the fact that the cortical regions that are selective for face
19
20 and word processing are located adjacent to the central visual information within the highest
21
22 level of retinotopic representation (Hasson et al., 2002; Levy et al., 2001) can be understood as
23
24 a natural consequence of topography and the constraint on local connectivity. Furthermore, the
25
26 cortical regions selective for these stimuli come to be located adjacent to retinotopic regions
27
28 coding foveal information, but in different hemispheres, with words in the left to permit
29
30 coordination with other language-related knowledge. Additionally, the VWFA and FFA are both
31
32 insensitive to low-level input variations (e.g. letter font; viewpoint for faces) suggesting that both
33
34 regions reflect functional specialization of higher-order visual cortex.
35
36
37

38 A further commonality is that lesions to each region renders the individual reliant on a
39
40 more piecemeal or segmental approach rather than one in which the configural or whole is
41
42 accessible. Just as vWFA lesions result in a laborious letter-by-letter sequential decoding of the
43
44 individual letters, lesions resulting in prosopagnosia give rise to a similar sequential process
45
46 with greater reliance on some features of the face (e.g. mouth, see example in Bukach et al,
47
48 2008) and a laborious encoding of the features as reflected in eye movement patterns (e.g., see
49
50 Stephan & Caine, 2009). Also, just as VWFA activation is affected by experience (Baker et al.,
51
52 2007; Wong et al., 2005), so adult-like face processing in the FFA also emerges with age and
53
54 experience and both regions evince a protracted developmental trajectory with signs of
55
56 specificity emerging roughly when children are in elementary school (Brem et al., 2005; Cohen
57
58
59
60

1
2
3 Kadosh & Johnson, 2007; Golarai et al., 2009, 2010; Joseph, Gathers & Bhatt, 2011; Scherf et
4
5 al. 2007)³. Finally, both regions are rather plastic: the VWFA can be acquired in the RH after left
6
7 occipital resection in childhood (Cohen, Lehericy et al., 2004) and there may also be a shift to
8
9 the right in patients following acquired LH lesions (Cohen, Lehericy et al., 2004; Cohen et al.,
10
11 2003). Relatedly, there are no apparent differences in the face recognition deficits of individuals
12
13 with unilateral lesions in infancy that impacted either the LH or RH (de Schonen et al., 2005).
14
15

16
17 One apparent challenge to our emphasis on the role of visual experience in shaping the
18
19 cortical organization of face and word processing is the recent observation by Reich et al.
20
21 (2011) that the VWFA, as localized in sighted individuals, is also the location of peak activation
22
23 in congenitally blind individuals reading Braille words (compared to nonsense Braille control
24
25 stimuli). It should be noted, however, that Reich et al. found reliable differences in the entire left
26
27 ventral occipitotemporal cortex all the way to V1, consistent with earlier findings by Noppeney,
28
29 Friston, & Price (2003) that were attributed to recruitment of these areas for semantic
30
31 processing. Reich et al.'s only evidence for (tactile) word-form representations per se was that
32
33 activation differences were larger for Braille reading compared with auditory verb generation,
34
35 but this comparison is possibly dubious as it involves separate and unrelated control tasks.
36
37 Moreover, the fact that activation differences peaked in the same location in blind and sighted
38
39 individuals may arise because of intrinsic patterns of connectivity between early visual cortex
40
41 and parietal structures involved in spatial attention (Greenberg et al., submitted manuscript) and
42
43 this inherent white matter arrangement biases where activation peaks are likely to be observed
44
45 in functional imaging studies.
46
47
48
49
50

51 **Where do hemispheric differences come from?**

- 52
53 3 The exact nature of the change over developmental time remains somewhat controversial with some
54
55 studies showing changes in the volume of activation for one category over another and others showing
56
57 a change in functional/effective connectivity over the course of development. The studies are also not
58
59 entirely consistent with each other (see Cantlon et al., 2010, showing adult-like activation to faces as
60
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.

1
2
3 Our account presupposes that the hemispheric differences and asymmetries emerge
4 over the course of experience, during which time the connectivity and topography constraints
5 play out. There are, however, other theoretical stances, which interpret the emergent
6 hemispheric differences as arising from a different origin. One such account is that the left
7 hemisphere (LH) processes input in an analytical or part-based way (hence its role with words)
8 whereas the right hemisphere (RH) processes input more holistically (hence its role with faces)
9 (see also Farah 1991, 1992). A second view attributes the word/face distinction to the distinction
10 between linguistic (LH) and spatial (RH) processing. Yet a further view is that the face/word
11 difference arises from differential frequency sensitivity (Robertson & Ivry, 2000) with the RH and
12 LH responding to relatively low and high spatial frequency information, respectively, and the
13 former critical for faces (RH), and the latter for words (LH). A final possibility is that the
14 face/word differences arise from the differential predisposition to process inputs categorically
15 (LH words) versus by coordinate relations (RH faces) (Kosslyn et al., 1989). These accounts all
16 suggest that it is the fundamental (perhaps even hard-wired or innately specified) properties of
17 the two hemispheres that play a role in shaping the underlying computational differences
18 between words and faces, but how and to what extent this is so remains to be determined.

19
20 Our approach is not mutually exclusive with these process-based accounts but it
21 emphasizes the importance of expertise in shaping cortical organization and function (Gauthier
22 & Bukach, 2007; Gauthier & Nelson, 2001), although expertise alone cannot explain why the
23 FFA and VWFA are located where they are, nor why other types of expertise (namely, those not
24 demanding high-acuity visual information) do not engage these areas. It also shares the
25 fundamental assumption of Malach and colleagues (Levy, Hasson, Hendler, & Malach, 2001;
26 Levy, Hasson, & Malach, 2004) regarding pressure for foveal acuity and cortical topography, but
27 goes beyond this by implementing the ideas in explicit simulations, enabling the testing of
28 specific predictions concerning the relationship of face and word processing. There remain of
29 course, complex questions about why the left hemisphere is language-dominant in the majority
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 of the population and the source of this organizational pattern. Such issues are beyond the
4
5 scope of this paper but are intriguing and remain to be addressed, too.
6
7

8 9 10 **Predictions**

11
12 A full assessment of the tractability of our account remains to be undertaken and, in particular,
13
14 there are a number of predictions that can be tested. Much of our work, thus far, has focused on
15
16 the need for representational cooperation between words and the language-related output, as
17
18 the key pressure that drives the left hemisphere specialization. But this cooperation occurs in
19
20 conjunction with representational competition between faces and words (given their
21
22 incompatibility as visual objects) and this competition, too, motivates the hemispheric
23
24 distinctions with left and right biased for words and faces, respectively. A prediction of this
25
26 trading relations view is that individuals who have greater asymmetries for faces (e.g. relative to
27
28 baseline, greater performance advantage for faces presented to left than right visual field or
29
30 greater activation in the right hemisphere in imaging) should show the converse for words on an
31
32 individual-by-individual basis, depending on how the cooperation and competition play out
33
34 during the course of development and experience. We also anticipate that some individuals will
35
36 have more bilaterally graded representation and that there will be a large range of individual
37
38 differences across the population. This prediction is eminently testable through half-field
39
40 studies, as well as functional imaging investigation, and we are currently undertaking such
41
42 explorations including examining hemispheric asymmetries for faces and words within individual
43
44 and across groups of young children, adolescents and young adults.
45
46
47
48

49 A further rather obvious prediction is that individuals with damage to the left VWFA and
50
51 presenting with pure alexia might also be impaired at face recognition, relative to normal
52
53 participants, albeit to a lesser extent than individuals with prosopagnosia following a lesion to
54
55 the right FFA. The converse is also predicted: individuals with a lesion to the right FFA and
56
57 presenting with prosopagnosia might also be impaired at word recognition, relative to controls,
58
59
60

1
2
3 albeit to a lesser extent than individuals with pure alexia following a lesion to the left VWFA. We
4
5 have examined these predictions in a small group of individuals, all of whom were pre-morbidly
6
7 normal and have acquired unilateral ventral cortex lesions (Behrmann & Plaut, in preparation).
8
9 In this study, we used the same series of face and of word experiments to evaluate the
10
11 performance of three adults with circumscribed unilateral right hemisphere lesions and
12
13 prosopagnosia and four pure alexic adults with circumscribed unilateral left hemisphere lesions.
14
15 Control participants matched to the two groups were also tested. In addition to the expected
16
17 impairment in face recognition, the prosopagnosic individuals showed abnormal word
18
19 recognition relative to the controls albeit not as marked as in the pure alexics and, in
20
21 complementary fashion, the pure alexic individuals showed abnormal face recognition relative to
22
23 the controls albeit to a lesser extent than in the prosopagnosics (for related findings, see
24
25 Buxbaum, Glosser, & Coslett, 1996, 1999). These empirical findings anchor a key prediction of
26
27 our account which is that hemispheric asymmetries for face and word recognition are graded
28
29 and not fully and independently segregated.
30
31
32
33
34
35

36 **Limitations and extensions**

37
38 The computational simulation presented here was intentionally kept as simple as possible in
39
40 order to provide the clearest illustration of the consequences of our putative computational
41
42 principles for graded specialization of the neural representations of faces and words (and, to a
43
44 more limited extent, houses). The most obvious limitations of the simulation are that it used a
45
46 small set of highly schematized stimuli, a strictly feedforward network architecture without lateral
47
48 and top-down interactions within or between hemispheres, a lack of separate excitatory and
49
50 inhibitory unit populations, and a biologically implausible learning procedure. Although we claim
51
52 that the core findings regarding learned functional specialization do not depend critically on any
53
54 of these simplifications, it is important to validate these findings in more realistic follow-up
55
56 simulations. Such follow-up versions should use realistic stimuli, employ a more biologically
57
58
59
60

1
2
3 plausible learning procedure (e.g., contrastive Hebbian learning; see O'Reilly, 1996), and permit
4
5 only excitatory connections between layers by using a separate population of local inhibitory
6
7 units within each layer. We do not anticipate that these elaborations will alter the basic
8
9 operation of the model but they will bring it into much closer alignment with the operation of real
10
11 neural systems.
12

13 14 15 16 **Conclusion**

17
18 Our central hypothesis is that the commonalities in the neural mechanisms of face and word
19
20 processing are not merely coincidental, as modular theories are left to conclude, but rather, are
21
22 the signature consequences of a set of general principles and constraints on neural computation
23
24 that operate throughout cortex. We note that these principles themselves are not intended to be
25
26 novel; in fact, we take them to be largely non-controversial. Instead, the novelty derives from
27
28 their common consequences for cortical organization and behavior in two seemingly unrelated
29
30 domains, specifically, in the context of words and faces,
31
32

33
34 Critically, when instantiated in explicit computational terms, these principles provide
35
36 insight into why each of these properties is partial rather than absolute. This is because the
37
38 principles and constraints are inherently graded—adherence to the forces of cooperation,
39
40 competition, and proximity in the process of learning cortical representations is a matter of
41
42 degree as these constraints trade off against each other, and thus the consequences for neural
43
44 and behavioral specialization is also graded. As a result, the implications of graded constraints
45
46 go beyond explaining why neither pure alexia nor prosopagnosia is entirely 'pure', and why,
47
48 across a host of imaging studies, the FFA and VWFA show substantial responses to stimuli
49
50 other than faces and words, respectively. They also imply that the functional and anatomical
51
52 division between face and word recognition should be graded—despite the clear differences
53
54 between the two domains, the FFA should be partially involved in word recognition and the
55
56 VWFA should be partially involved in face recognition. In this way, our theoretical perspective
57
58
59
60

1
2
3 leads to important and otherwise unexpected predictions concerning the partial co-mingling of
4
5 face and word processing, including face recognition impairments in pure alexia, word
6
7 recognition impairments in prosopagnosia, graded participation of the FFA and VWFA in
8
9 face/word recognition in normal observers, and a number of other implications that remain to be
10
11 tested.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure legends

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, 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 co-ordinates—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 grey-scale 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 favoring short connections.

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.

Figure 5. Example receptive (input-to-intermediate) and projective (intermediate-to-output) fields for two units in the right-hemisphere intermediate layer. (Units in the left-hemisphere intermediate layer show essentially the same patterns of receptive and projective fields as those shown here.) Each square in the receptive and projective fields shows the value of the weight (red for positive weights; blue for negative weights) from that unit either into or out of the depicted unit in the intermediate layer (outlined in yellow). 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 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).

References

- 1
2
3
4
5 Allison, T., McCarthy, G., Nobre, A.C., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and
6 the perception of faces, words, numbers and colors. *Cerebral Cortex*, *5*, 544-554.
- 7 Avidan, G., & Behrmann, M. (2009). Functional MRI reveals compromised neural integrity of the face
8 processing network in congenital prosopagnosia *Current Biology*, *19*, 1146-1150.
- 9 Bach, F.R., & Jordan, M.I. (2002). Kernel independent component analysis. *Journal of Machine Learning*
10 *Research*, *3*, 1-48.
- 11 Baker, C.I., Liu, J., Wald, L.L., Kwong, K., Benner, T., & Kanwisher, N. (2007). Visual word processing
12 and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the*
13 *National Academy of Science, USA*, *104*, 9087-9092.
- 14 Bartlett, M.S., Movellan, J.R., & Sejnowski, T.J. (2002). Face recognition by independent component
15 analysis. *IEEE Transactions on Neural Networks*, *13*, 1450-1464.
- 16 Barton, J.J., Cherkasova, M.V., Press, D.Z., Intriligator, J.M., & O'Connor, M. (2004). Perceptual functions
17 in prosopagnosia. *Perception*, *33*, 939-956.
- 18 Barton, J.J.S. (2008). Structure and function in acquired prosopagnosia: lessons from a series of 10
19 patients with brain damage. *Journal of Neuropsychology*, *2*, 197-225.
- 20 Barton, J.J.S. (2009). What is meant by impaired configural processing in acquired prosopagnosia?
21 *Perception*, *38*, 242-260.
- 22 Behrmann, M., & Kimchi, R. (2003). What does visual agnosia tell us about perceptual organization and
23 its relationship to object perception? *Journal of Experimental Psychology: Human Perception and*
24 *Performance*, *29*, 19-42.
- 25 Behrmann, M., Nelson, J., & Sekuler, E. (1998). Visual complexity in letter-by-letter reading: "Pure" alexia
26 is not so pure. *Neuropsychologia*, *36*, 1115-1132.
- 27 Behrmann, M., & Plaut, D.C. (in preparation). *Faces and words: Flip sides of the same brain.*
- 28 Bilalic, M., Langner, R., Ulrich, R., & Grodd, W. (2011). Many faces of expertise: fusiform face area in
29 chess experts and novices. *Journal of Neuroscience*, *31*, 10206-10214.
- 30 Blanz, V., & Vetter, T. (2003). Face recognition based on fitting a 3D morphable model. *IEEE Transac-*
31 *tions on Pattern Analysis and Machine Intelligence*, *25*, 1063-1074.
- 32 Bouvier, S.E., & Engel, S.A. (2006). Behavioral deficits and cortical damage loci in cerebral
33 achromatopsia. *Cerebral Cortex*, *16*, 183-191.
- 34 Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., & Brandeis, D. (2006). Evidence
35 for developmental changes in the visual word processing network beyond adolescence.
36 *Neuroimage*, *29*, 822-837.
- 37 Büchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway.
38 *Nature*, *394*, 274-277.
- 39 Bukach, C., Le Grand, R., Kaiser, M.D., Bub, D.N., & Tanaka, J.W. (2008). Preservation of mouth region
40 processing in two cases of prosopagnosia. *Journal of Neuropsychology*, *2*, 227-244.
- 41 Burgund, E.D., Schlaggar, B.L., & Petersen, S.E. (2006). Development of letter-specific processing: The
42 effect of reading ability. *Acta Psychologica (Amsterdam)*. *122*, 99-108.
- 43 Burton, M.A., Bruce, V., & Hancock, P.J.B. (1999). From pixels to people: A model of familiar face
44 recognition. *Cognitive Science*, *23*, 1-31.
- 45 Burton, M.A., Young, A.W., Bruce, V., Johnston, R.A., & Ellis, A.W. (1991). Understanding covert
46 recognition. *Cognition*, *39*, 129-166.
- 47 Busigny, T., & Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*,
48 *46*, 965-981.
- 49 Buxbaum, L., Glosser, G., & Coslett, H.B. (1996). Relative sparing of object recognition in alexia-
50 prosopagnosia. *Brain and Cognition*, *32*, 202-205.
- 51 Buxbaum, L.J., Glosser, G., & Coslett, H.B. (1999). Impaired face and word recognition without object
52 agnosia. *Neuropsychologia*, *37*, 41-50.
- 53 Cai, Q., Lavidor, M., Brysbaert, M., Paulignan, Y., & Nazir, T.A. (2008). Cerebral lateralization of frontal
54 lobe language processes and lateralization of the posterior visual word processing system.
55 *Journal of Cognitive Neuroscience*, *20*, 672-681.
- 56 Cantlon, J.F., Pinel, P., Dehaene, S., & Pelphrey, K.A. (2011) Cortical representations of symbols, ob-
57 jects, and faces Are Pruned Back during Early Childhood. *Cerebral Cortex*, *21*, 191-199.
- 58 Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: the case for the visual word
59
60

- 1
2
3 form area. *Neuroimage*, 22(1), 466-476.
- 4 Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M.A., et al. (2000).
5 The visual word form area: spatial and temporal characterization of an initial stage of reading in
6 normal subjects and posterior split-brain patients. *Brain*, 123, 291-307.
- 7 Cohen, L., Henry, C., Dehaene, S., Martinaud, O., Lehericy, S., Lemer, C., et al. (2004). The
8 pathophysiology of letter-by-letter reading. *Neuropsychologia*, 42, 1768-1780.
- 9 Cohen, L., Lehericy, S., Henry, C., Bourgeois, M., Larroque, C., Sainte-Rose, C., et al. (2004). Learning
10 to read without a left occipital lobe: right-hemispheric shift of visual word form area. *Annals of*
11 *Neurology*, 56, 890-894.
- 12 Cohen, L., Martinaud, O., Lemer, C., Lehericy, S., Samson, Y., Obadia, M., et al. (2003). Visual word
13 recognition in the left and right hemispheres: Anatomical and functional correlates of peripheral
14 alexias. *Cerebral Cortex*, 13, 1313-1333.
- 15 Cohen Kadosh, K., Cohen Kadosh, R., Dick, F., & Johnson, M.H. (2010). Developmental Changes in Ef-
16 fective Connectivity in the Emerging Core Face Network. *Cerebral Cortex*, 21, 1389-1394.
- 17 Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model
18 of visual word recognition and reading aloud. *Psychological Review*, 108, 204-256.
- 19 Cowey, A. (1979). Cortical maps and visual perception. The Grindley Memorial Lecture. *Quarterly*
20 *Journal of Experimental Psychology*, 31, 1-17.
- 21 Dailey, M.N., & Cottrell, G.W. (1999). Organization of face and object recognition in modular neural
22 networks. *Neural Networks*, 12, 1053-1074.
- 23 Dailey, M.N., Cottrell, G.W., Padgett, C., & Adolphs, R. (2002) EMPATH: A neural network that
24 categorizes facial expressions. *Journal of Cognitive Neuroscience* 14, 1158-1173.
- 25 Damasio, A., Damasio, H., & Tranel, D. (1986). Prosopagnosia: Anatomic and physiological aspects. In
26 H.D. Ellis, M.A. Jeeves, F. Newcombe & A. Young (Eds.), *Aspects of face processing* (pp. 279-
27 290). Dordrecht: Martinus Nijhoff.
- 28 Davis, C.J. (1999). *The Self-Organizing Lexical Acquisition and Recognition (SOLAR) Model of Visual*
29 *Word Recognition*. Unpublished Doctoral Dissertation. University of New South Wales.
- 30 De Yoe, E.A., Carman, G.J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., et al. (1996). Mapping
31 striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National*
32 *Academy of Sciences*, 93, 2382-2386.
- 33 Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56, 384-398.
- 34 Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in*
35 *Cognitive Science*, 15, 254-262.
- 36 Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal.
37 *Trends in Cognitive Science*, 9, 335-341.
- 38 Dehaene, S., Pegado, F., Braga, L.W., Ventura, P., Filho, G.N., Jobert, A., et al. (2010). How Learning to
39 Read Changes the Cortical Networks for Vision and Language. *Science*, 330, 1359-1364.
- 40 Devlin, J.T., Jamison, H.L., Gonnerman, L.M., & Matthews, P.M. (2006). The role of the posterior fusiform
41 gyrus in reading. *Journal of Cognitive Neuroscience*, 18, 911-922.
- 42 Downing, P.E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual
43 processing of the human body. *Science*, 293, 2470-2473.
- 44 Epstein, R.A., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area:
45 recognition, navigation, or encoding? *Neuron*, 23, 115-125.
- 46 Epstein, R.A., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*,
47 392, 598-601.
- 48 Epstein, R.A., Deyoe, E.A., Press, D.Z., Rosen, A.C., & Kanwisher, N. (2001). Neuropsychological evi-
49 dence for a topographical learning mechanism in parahippocampal cortex. *Cognitive Neuropsy-*
50 *chology*, 18, 481-508.
- 51 Etemad, K., & Chellappa, R. (1997). Discriminant analysis for recognition of human face images. *Journal*
52 *of the Optical Society of America A*, 14, 1724-1733.
- 53 Farah, M.J. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual
54 object recognition. *Cognitive Neuropsychology*, 8, 1-19.
- 55 Farah, M.J. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of
56 domain specificity in visual object recognition. *Current Directions in Psychological Science*, 1,
57 164-169.
- 58 Farah, M.J. (1996). Is face recognition 'special'? Evidence from neuropsychology. *Behavioural Brain*
59
60

- 1
2
3 *Research*, 76, 181-189.
- 4 Farah, M.J. (1997). "Visual object agnosia without alexia or prosopagnosia: Arguments for separate
5 knowledge stores" : Reply. *Visual Cognition*, 4, 219-220.
- 6 Farah, M.J. (1999). Relations among the agnosias. In G.W. Humphreys (Ed.), *Case studies in the*
7 *neuropsychology of vision* (pp. 181-200). Hove, UK: Psychology Press.
- 8 Farah, M.J. (Ed.) (2004). *Visual agnosia*. Cambridge, MA: MIT Press.
- 9 Farah, M.J., O'Reilly, R.C., & Vecera, S.P. (1993). Dissociated overt and covert recognition as an
10 emergent property of a lesioned neural network. *Psychological Review*, 100, 571-588.
- 11 Farah, M.J., Tanaka, J.W., & Drain, H.M. (1995). What causes the face inversion effect? *Journal of*
12 *Experimental Psychology: Human Perception and Performance*, 21, 628-634.
- 13 Feinberg, T.E., Schindler, R.J., Ochoa, E., Kwan, P.C., & Farah, M.J. (1994). Associative visual agnosia
14 and alexia without prosopagnosia. *Cortex*, 30, 395-412.
- 15 Fiez, J.A., Balota, D.A., Raichle, M.E., & Petersen, S.E. (1999). Effects of lexicality, frequency, and
16 spelling-to-sound consistency on the functional anatomy of reading. *Neuron*, 24, 205-218.
- 17 Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- 18 Friedman, R.B., & Alexander, M.P. (1984). Pictures, images, and pure alexia: A case study. *Cognitive*
19 *Neuropsychology*, 1, 9-23.
- 20 Furl, N., Garrido, L., Dolan, R.J., Driver, J., & Duchaine, B. (2010). Fusiform gyrus face selectivity relates
21 to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience*, 23, 1723-
22 1740.
- 23 Gainotti, G., & Marra, (2011). Differential contribution of right and left temporo-occipital and anterior tem-
24 poral lesions to face recognition disorders, *Frontiers in Human Neuroscience*, 5, 1-11.
- 25 Gauthier, I., Tarr, M.J., Anderson, A.W., Skudlarski, P., & Gore, J.C. (1999). Activation of the middle
26 fusiform "face area" increases with expertise in recognizing novel objects. *Nature Neuroscience*,
27 2, 568-573.
- 28 Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C., & Anderson, A.W. (2000). The fusiform
29 "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive*
30 *Neuroscience*, 12, 495-504.
- 31 Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition*, 103, 322-330.
- 32 Gauthier, I., & Nelson, C.A. (2001). The development of face expertise. *Current Opinion in Neurobiology*,
33 11, 219-24.
- 34 Golarai, G., Ghahremani, D.G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J.L., Gabrieli, J.D., & Grill-
35 Spector, K. (2007) Differential development of high-level visual cortex correlates with cate-
36 gory-specific recognition memory. *Nature Neuroscience*, 10, 512-522.
- 37 Golarai, G., Liberman, A., Yoon, J.M., & Grill-Spector, K. (2010). Differential development of the ventral
38 visual cortex extends through adolescence. *Frontiers in Human Neuroscience*, 3, 80.
- 39 Grainger J., Granier J.P., Farioli F., Van Assche E., & van Heuven W. (2006). Letter position information
40 and printed word perception: The relative-position priming constraint. *Journal of Experimental*
41 *Psychology: Human Perception and Performance*, 32, 865-884.
- 42 Greenberg, A., Verstynen, T., Chiu, Y-V., Yantis, S., Schneider, W., & Behrmann, M. (under review).
43 *Spatiotopic structural connectivity underlying visual attention*.
- 44 Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, 27,
45 649-677.
- 46 Grill-Spector, K., Sayres, R., & Ress, D. (2006). High-resolution imaging reveals highly selective non-face
47 clusters in the fusiform face area. *Nature Neuroscience*, 9, 1177-1185.
- 48 Hanson, S.J., & Schmidt, A. (2011). High-resolution imaging of the fusiform face area (FFA) using
49 multivariate non-linear classifiers shows diagnosticity for non-face categories. *Neuroimage*, 54,
50 1715-1734.
- 51 Harm, M.W., & Seidenberg, M.S. (1999). Phonology, reading acquisition, and dyslexia: Insights from
52 connectionist models. *Psychological Review*, 106, 491-528.
- 53 Harm, M.W., & Seidenberg, M.S. (2004). Computing the meanings of words in reading: Cooperative
54 division of labor between visual and phonological processes. *Psychological Review*, 111, 662-
55 720.
- 56 Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Center-biased representation for
57 characters in the human ventral visual stream. *Neuron*, 34, 479-490.
- 58 Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., & Pietrini, P. (2001). Distributed and
59
60

- 1
2
3 overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425-
4 2429.
- 5 Haxby, J.V. (2006). Fine structure in representations of faces and objects. *Nature Neuroscience*, 9, 1084-
6 1086.
- 7 Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., & Pietrini, P. (2001). Distributed and
8 overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425-
9 2429.
- 10 Haxby, J.V., Petit, L., Ungerleider, L.G., & Courtney, S.M. (2000). Distinguishing the functional roles of
11 multiple regions in distributed neural systems for visual working memory. *Neuroimage*, 11, 380-
12 391.
- 13 Hoffman, E.A., & Haxby, J.V. (2000). Distinct representations of eye gaze and identity in the distributed
14 human neural system for face perception. *Nature Neuroscience*, 3, 80-84.
- 15 Hsiao, J., Shieh, D., & Cottrell, G.W. (2008). Convergence of the visual field split: Hemispheric modeling
16 of face and object recognition. *Journal of Cognitive Neuroscience*, 20, 2298-2307.
- 17 Ishai, A. (2008). Let's face it: it's a cortical network. *Neuroimage*, 40, 415-419.
- 18 Ishai, A., Schmidt, C.F., & Boesiger, P. (2005). Face perception is mediated by a distributed cortical
19 network. *Brain Research Bulletin*, 67, 87-93.
- 20 Jacobs, R.A. (1997). Nature, nurture and the development of functional specializations: A computational
21 approach. *Psychonomic Bulletin and Review*, 4, 299-309.
- 22 Jacobs, R.A., & Jordan, M.I. (1992). Computational Consequences of a Bias toward Short Connections.
23 *Journal of Cognitive Neuroscience*, 4, 323-336.
- 24 Joseph, J.E. Gathers, A.D., & Bhatt, R.S. (2011). Progressive and regressive developmental changes in
25 neural substrates for face processing: Testing specific predictions of the Interactive Specialization
26 account. *Developmental Science*, 14, 227-241.
- 27 Kanan, C.M. & Cottrell, G.W. (2010). Robust classification of objects, faces, and flowers using natural
28 image statistics. In *Proceedings of the Computer Vision and Pattern Recognition Conference*
29 *(CVPR-2010)*.
- 30 Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, 3, 759-763.
- 31 Kanwisher, N. (2010). Functional specificity in the human brain: a window into the functional architecture
32 of the mind. *Proceedings of the National Academy of Science, USA*, 107, 11163-11170.
- 33 Kanwisher, N., McDermott, J., & Chun, M.M. (1997). The fusiform face area: A module in human
34 extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17, 4302-4311.
- 35 Kanwisher, N., Woods, R.P., Iacoboni, M., & Mazziotta, J.C. (1997). A locus in human extrastriate cortex
36 for visual shape analysis. *Journal of Cognitive Neuroscience*, 9, 133-142.
- 37 Kanwisher, N.G., McDermott, J., & Chun, M.M. (1997). The fusiform face area: A module in human
38 extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17, 4302-4311.
- 39 Kleinschmidt, A., & Cohen, L. (2006). The neural bases of prosopagnosia and pure alexia: recent insights
40 from functional neuroimaging. *Current Opinion in Neurology*, 19, 386-391.
- 41 Koriat, A., & Norman, J. (1989). Why is word recognition impaired by disorientation while the identification
42 of single letters is not? *Journal of Experimental Psychology: Human Perception and*
43 *Performance*, 15, 153-163.
- 44 Kosslyn, S.M., Koenig, O., Barrett, A., Cave, C.B., Tang, J., & Gabrieli, J.D.E. (1989). Evidence for two
45 types of spatial representations: Hemispheric specialization for categorical and coordinate
46 relations. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 723-735.
- 47 Kriegeskorte, N., Formisano, E., Sorge, B., & Goebel, R. (2007). Individual faces elicit distinct response
48 patterns in human anterior temporal cortex. *Proceedings of the National Academy of Science,*
49 *USA*, 104, 20600-20605.
- 50 Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word
51 form area and the frequency with which words are encountered: evidence from a parametric fMRI
52 study. *Neuroimage*, 21, 946-953.
- 53 Lavidor, M., Ellis, A.W., Shillcock, R., & Bland, T. (2001). Evaluating a split processing model of visual
54 word recognition: effects of word length. *Cognitive Brain Research*, 12, 265-272.
- 55 Leder, H., & Bruce, V. (2000). When inverted faces are recognized: the role of configural information in
56 face recognition. *Quarterly Journal of Experimental Psychology A*, 53, 513-536.
- 57 Levy, I., Hasson, U., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object
58 areas. *Nature Neuroscience*, 4, 533-539.
- 59
60

- 1
2
3 Levy, I., Hasson, U., & Malach, R. (2004). One picture is worth at least a million neurons. *Current Biology*,
4 14, 996-1001.
- 5 Liu, Y.C., Wang, A.G., & Yen, M.Y. "Seeing but not identifying": pure alexia coincident with prosopagnosia
6 in occipital arteriovenous malformation, *Graefe's Archive of Clinical and Experimental Ophthalmology*,
7 248, 1087-1089.
- 8 Mahon, B.Z. & Caramazza, A. (2011). What drives the organization of knowledge in the brain. *Trends in*
9 *Cognitive Sciences*, 15, 97-103.
- 10 Martelli, M., Majaj, N.J., & Pelli, D.G. (2005). Are faces processed like words? A diagnostic test for
11 recognition by parts. *Journal of Vision*, 5, 58-70.
- 12 Maurer, D., Le Grand, R., & Mondloch, C.J. (2002). The many faces of configural processing. *Trends in*
13 *Cognitive Sciences*, 6, 255-260.
- 14 Maurer, U., Brandeis, D., & McCandliss, B.D. (2005). Fast, visual specialization for reading in English
15 revealed by the topography of the N170 ERP response. *Behavioral and Brain Functions*, 1, 13.
- 16 McCandliss, B.D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in
17 the fusiform gyrus. *Trends in Cognitive Science*, 7, 293-299.
- 18 McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter
19 perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375-407.
- 20 McClelland, J.L., & Rumelhart, D.E. (1985). Distributed memory and the representation of general and
21 specific information. *Journal of Experimental Psychology: General*, 114, 159-188.
- 22 McKone, E. & Robbins, R. (in press). Are faces special? In A.J. Calder, G. Rhodes, J.V. Haxby and M.H.
23 Johnson (Eds.), *The handbook of face perception*. Oxford, UK: Oxford University Press.
- 24 Mechelli, A., Gorno-Tempini, M.L., & Price, C.J. (2003). Neuroimaging studies of word and pseudoword
25 reading: consistencies, inconsistencies, and limitations. *Journal of Cognitive Neuroscience*, 15,
26 260-271.
- 27 Moghaddam, B., Jebara, T., & Pentland, A. (2000). Bayesian face recognition. *Pattern Recognition*, 33,
28 1771-1782.
- 29 Montant, M., & Behrmann, M. (2000). Pure Alexia: review of cases. *Neurocase*, 6, 265-294.
- 30 Mozer, M.C. (1991). *The perception of multiple objects: A connectionist approach*. Cambridge, MA: MIT
31 Press.
- 32 Mozer, M.C., & Behrmann, M. (1990). On the interaction of selective attention and lexical knowledge: A
33 connectionist account of neglect dyslexia. *Journal of Cognitive Neuroscience*, 2, 96-123.
- 34 Murtha, S., Chertkow, H., Beaugregard, M., & Evans, A. (1999). The neural substrate of picture naming.
35 *Journal of Cognitive Neuroscience*, 11, 399-423.
- 36 Nelson, M.E., & Bower, J.M. (1990). Brain maps and parallel computers. *Trends in Neurosciences*, 13,
37 403-408.
- 38 Nestor, A., Plaut, D.C. & Behrmann, M. (2011). Unraveling the distributed neural code of facial identity
39 through spatiotemporal pattern analysis, ePub ahead of print.
- 40 Nestor, A., Behrmann, M. & Plaut, D.C. (submitted). *The neural basis of visual word form processing – a*
41 *multivariate investigation*.
- 42 Noppeney, U., Friston, K.J., & Price, C.J. (2003). Effects of visual deprivation on the organization of the
43 semantic system. *Brain*, 126, 1620–1627.
- 44 Norman, K.A., Polyn, S.M., Detre, G.J., & Haxby, J.V. (2006). Beyond mind-reading: multi-voxel pattern
45 analysis of fMRI data. *Trends in Cognitive Science*, 10, 424-430.
- 46 O'Reilly, R.C. (1996). Biologically plausible error-driven learning using local activation differences: The
47 generalized recirculation algorithm. *Neural Computation*, 8, 895-938.
- 48 Peelen, M.V., & Downing, P.E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of*
49 *Neurophysiology*, 93, 603-608.
- 50 Perry, C., Ziegler, J.C., & Zorzi, M. (2007). Nested incremental modeling in the development of
51 computational theories: The CDP+ model of reading aloud. *Psychological Review*, 114, 273-315.
- 52 Petersen, S., Fox, P.T., Snyder, A., & Raichle, M.E. (1990). Activation of extrastriate and frontal cortical
53 areas by visual words and word-like stimuli. *Science*, 249, 1041-1044.
- 54 Petersen, S.E., & Fiez, J. (1993). The processing of single words studies with positron emission
55 tomography. *Annual Review of Neurosciences*, 16, 509-530.
- 56 Plaut, D.C. (1999). A connectionist approach to word reading and acquired dyslexia: Extension to
57 sequential processing. *Cognitive Science*, 23, 543-568.
- 58 Plaut, D.C. (2002). Graded modality-specific specialization in semantics: A computational account of optic
59
60

- 1
2
3 aphasia. *Cognitive Neuropsychology*, 19, 603-639.
- 4 Plaut, D.C., McClelland, J.L., Seidenberg M.S., & Patterson, K. (1996) Understanding normal and
5 impaired word reading: Computational principles in quasi-regular domains. *Psychological*
6 *Review*, 103, 56-115.
- 7 Polk, T.A., & Farah, M.J. (2002). Functional MRI evidence for an abstract, not perceptual, word-form area.
8 *Journal of Experimental Psychology: General*, 131, 65-72.
- 9 Polk, T.A., Stallcup, M., Aguirre, G.K., Alsop, D.C., D'Esposito, M., Detre, J.A., et al. (2002). Neural
10 specialization for letter recognition. *Journal of Cognitive Neuroscience*, 14, 145-159.
- 11 Price, C.J., & Devlin, J.T. (2003). The myth of the visual word form area. *Neuroimage*, 19, 473-481.
- 12 Price, C.J., & Devlin, J.T. (2011). The Interactive Account of ventral occipitotemporal contributions to
13 reading. *Trends in Cognitive Science*, in press.
- 14 Price, C.J., & Mechelli, A. (2005). Reading and reading disturbance. *Current Opinion in Neurobiology*, 15,
15 231-238.
- 16 Puce, A., Allison, T., Asgari, M., Gore, J.C., & McCarthy, G. (1996). Differential sensitivity of human visual
17 cortex to faces, letter strings, and textures: A functional magnetic resonance imaging study. *The*
18 *Journal of Neuroscience*, 16, 5205-5215.
- 19 Puce, A., Allison, T., Gore, J.C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate
20 cortex studied by functional MRI. *Journal of Neurophysiology*, 74, 1192-1199.
- 21 Quartz, S.R. & Sejnowski, T.J. (1997). The neural basis of cognitive development: A constructivist
22 manifesto. *Behavioral and Brain Sciences*, 20, 537-596.
- 23 Rajimehr, R., Young, J.C., & Tootell, R.B. (2009). An anterior temporal face patch in human cortex,
24 predicted by macaque maps. *Proceedings of the National Academy of Science, USA*, 106, 1995-
25 2000.
- 26 Reich, L, Szwed, M., Cohen, L., & Amedi, A. (2011). A ventral visual stream reading center independent
27 of visual experience. *Current Biology*, 21, 363-368.
- 28 Reddy, L., & Kanwisher, N. (2006). Coding of visual objects in the ventral stream. *Current Opinion in*
29 *Neurobiology*, 16, 408-414.
- 30 Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1986). Learning representations by back- propagating
31 errors. *Nature*, 323, 533-536.
- 32 Rumelhart, D.E., & McClelland, J.L. (1982). An interactive activation model of context effects in letter
33 perception: Part 2. The contextual enhancement effect and some tests and extensions of the
34 model. *Psychological Review*, 89, 60-94.
- 35 Salvan, C.V., Ulmer, J.L., DeYoe, E.A., Wascher, T., Mathews, V.P., Lewis, J.W., et al. (2004). Visual
36 object agnosia and pure word alexia: correlation of functional magnetic resonance imaging and
37 lesion localization. *Journal of Computer Assisted Tomography*, 28, 63-67.
- 38 Scherf, K.S., Behrmann, M., Humphreys, K., & Luna, B. (2007). Visual category-selectivity for faces,
39 places and objects emerges along different developmental trajectories. *Developmental Science*,
40 10, 15-30.
- 41 Schwarzlose, R.F., Baker, C.I., & Kanwisher, N. (2005). Separate face and body selectivity on the
42 fusiform gyrus. *Journal of Neuroscience*, 25, 11055-11059.
- 43 Sekuler, E.B., & Behrmann, M. (1996). Perceptual cues in pure alexia. *Cognitive Neuropsychology*, 13,
44 941-974.
- 45 Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., et al. (1995). Borders of
46 multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*,
47 268, 889-893.
- 48 Sergent, J., Ohta, S., & MacDonald, B. (1992a). Functional neuroanatomy of face and object processing.
49 *Brain*, 115, 15-36.
- 50 Sergent, J., Ohta, S., & MacDonald, B. (1992b). Neuroanatomy of face and object processing. *Brain*, 115,
51 15-36.
- 52 Sowards, T.V. (2011). Neural structures and mechanisms involved in scene recognition: A review and
53 interpretation. *Neuropsychologia*, 49, 3, 277-298.
- 54 Shillcock, R., Ellison, M.T., & Monaghan, P. (2000). Eye-fixation behavior, lexical storage and visual word
55 recognition in a split processing model. *Psychological Review*, 107, 824-851.
- 56 Spiridon, M., Fischl, B., & Kanwisher, N. (2006). Location and spatial profile of category-specific regions
57 in human extrastriate cortex. *Human Brain Mapping*, 27(1), 77-89.
- 58 Spiridon, M., & Kanwisher, N. (2002). How distributed is visual category information in human occipito-
59
60

- 1
2
3 temporal cortex? An fMRI study. *Neuron*, 35(6), 1157-1165.
- 4 Starrfelt, R., & Behrmann, M. (2011). Number reading in pure alexia: A review. *Neuropsychologia*, 49, 9,
5 2283-2298.
- 6 Stephan, B.C., & Caine, D. (2009). Aberrant pattern of scanning in prosopagnosia reflects impaired face
7 processing. *Brain and Cognition*, 69, 262-268.
- 8 Tagamets, M.A., Novick, J.M., Chalmers, M.L., & Friedman, R.B. (2000). A parametric approach to
9 orthographic processing in the brain: An fMRI study. *Journal of Cognitive Neuroscience*, 12, 281-
10 297.
- 11 Tanaka, J.W., & Farah, M.J. (1993). Parts and wholes in face recognition. *Quarterly Journal of*
12 *Experimental Psychology*, 46A, 225-245.
- 13 Taylor, J.C., Wiggett, A.J., & Downing, P.E. (2010). fMRI-adaptation studies of viewpoint tuning in the
14 extrastriate and fusiform body areas. *Journal of Neurophysiology*, 103, 1467-77.
- 15 Tong, M.H., Joyce, C.A., & Cottrell, G.W. (2008). Why is the fusiform face area recruited for novel
16 categories of expertise? A neurocomputational investigation. *Brain Research*, 1202, 14-24.
- 17 Tootell, R.B.H., Mendola, J., Hadjikhani, H., Ledden, P., Liu, A., Reppas, J., et al. (1997). Functional
18 analysis of V3a and related areas in human visual cortex. *Journal of Neuroscience*, 71, 7060-
19 7078.
- 20 Tranel, D., Damasio, H., & Damasio, A.R. (1997). A neural basis for the retrieval of conceptual
21 knowledge. *Neuropsychologia*, 35, 1319-1328.
- 22 Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3, 71-86.
- 23 Turkeltaub, P.E., Eden, G.F., Jones, K.M., & Zeffiro, T.A. (2002). Meta-analysis of the functional
24 neuroanatomy of single-word reading: method and validation. *Neuroimage*, 16, 765-780.
- 25 Warrington, E.K., & Shallice, T. (1980). Word-form dyslexia. *Brain*, 103, 99-112.
- 26 Weekes, B.S. (1997). Differential effects of number of letters on word and nonword naming latency.
27 *Quarterly Journal of Experimental Psychology*, 50A, 439-456.
- 28 Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and
29 selective literature review. *Psychonomic Bulletin and Review*, 8, 221-243.
- 30 Willems, R.M., Peelen, M.V., & Hagoort, P. (2010). Cerebral lateralization of face-selective and body-
31 selective visual areas depends on handedness. *Cerebral Cortex*, 20, 1719-1725.
- 32 Yang, M.-H. (2002). Face recognition using kernel methods. In T. Diederich, S. Becker, Z. Ghahramani
33 (Eds.), *Advances in Neural Information Processing Systems, Volume 14* (pp. 1457-1464).
34 Cambridge, MA: MIT Press.
- 35 Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current*
36 *Biology*, 15, 2256-2262.
- 37 Yovel, G., Tambini, A., & Brandman, T. (2008). The asymmetry of the fusiform face area is a stable
38 individual characteristic that underlies the left-visual-field superiority for faces. *Neuropsychologia*,
39 46, 3061-3068.
- 40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1a

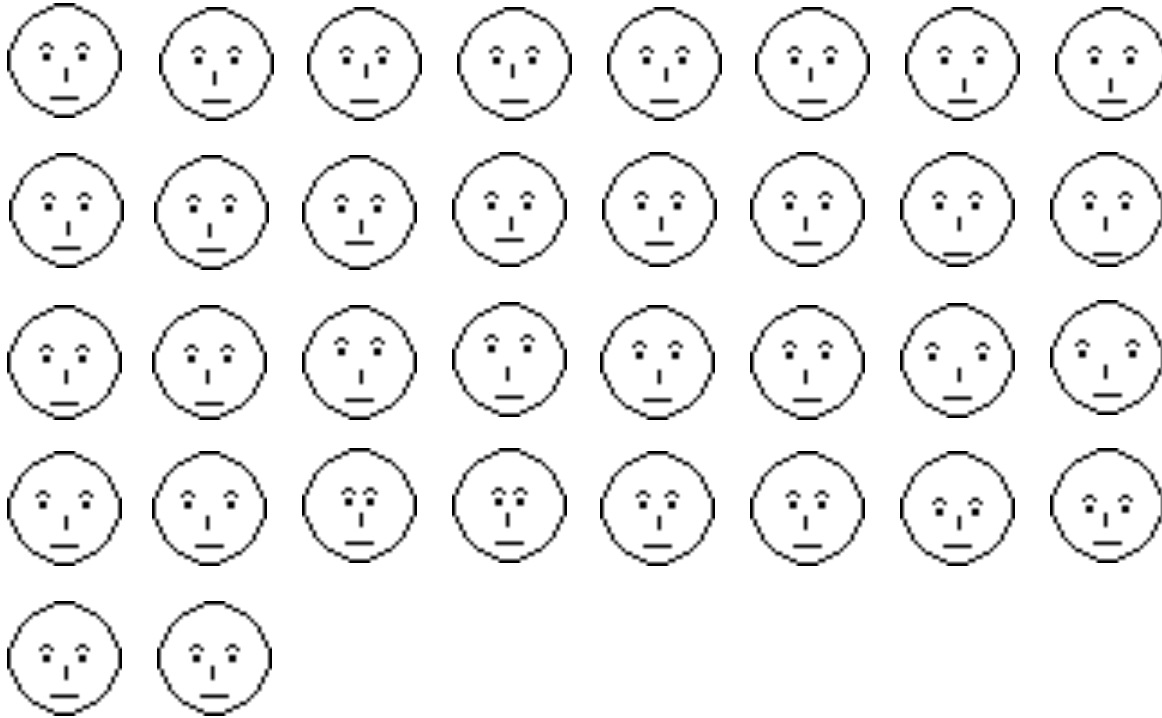


Figure 1b

BAG	BED	BEG	BET	BID	BIG	BIT	BUD
BUT	DAD	DID	DIG	DIP	DOG	DOT	DUG
MUD	MET	MEN	MAT	MAP	MAN	MAD	PAD
PAN	PEN	PET	PIG	PIP	POP	POT	PUT
TAG	TAN	TAP	TEN	TIN	TIP	TON	TOP

Figure 1c



Figure 2

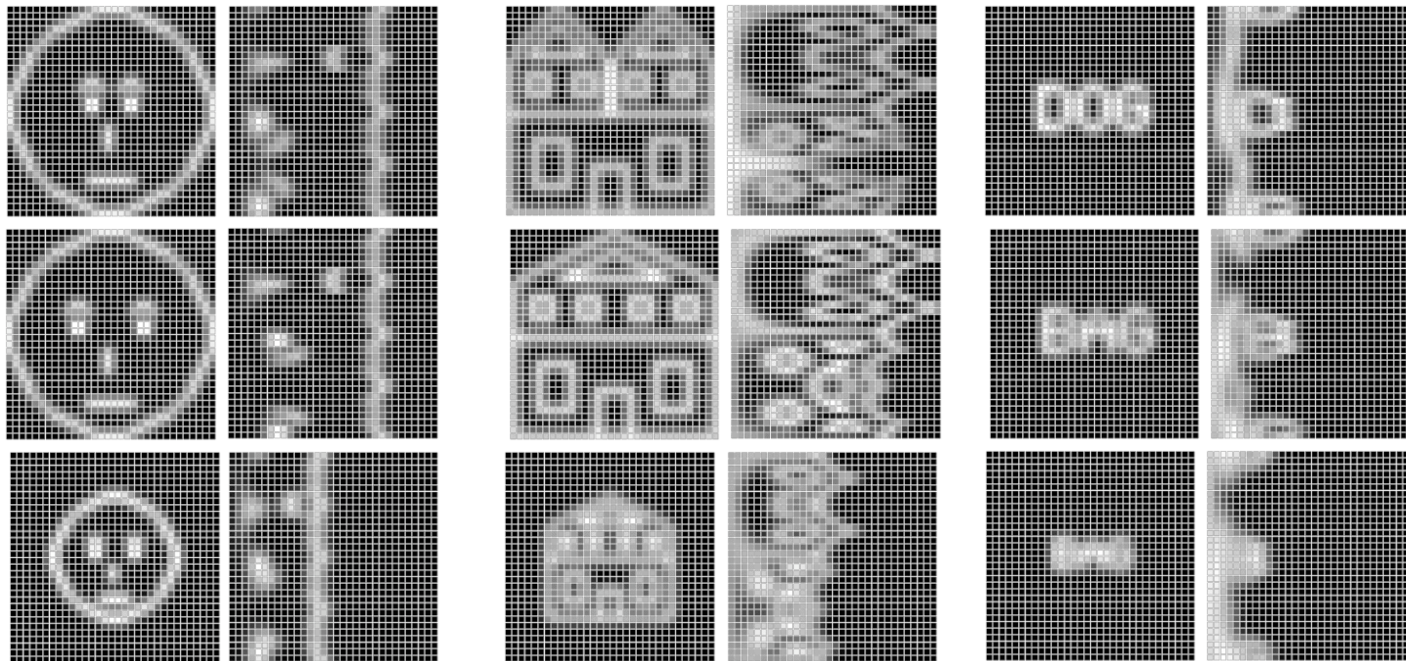


Figure 3

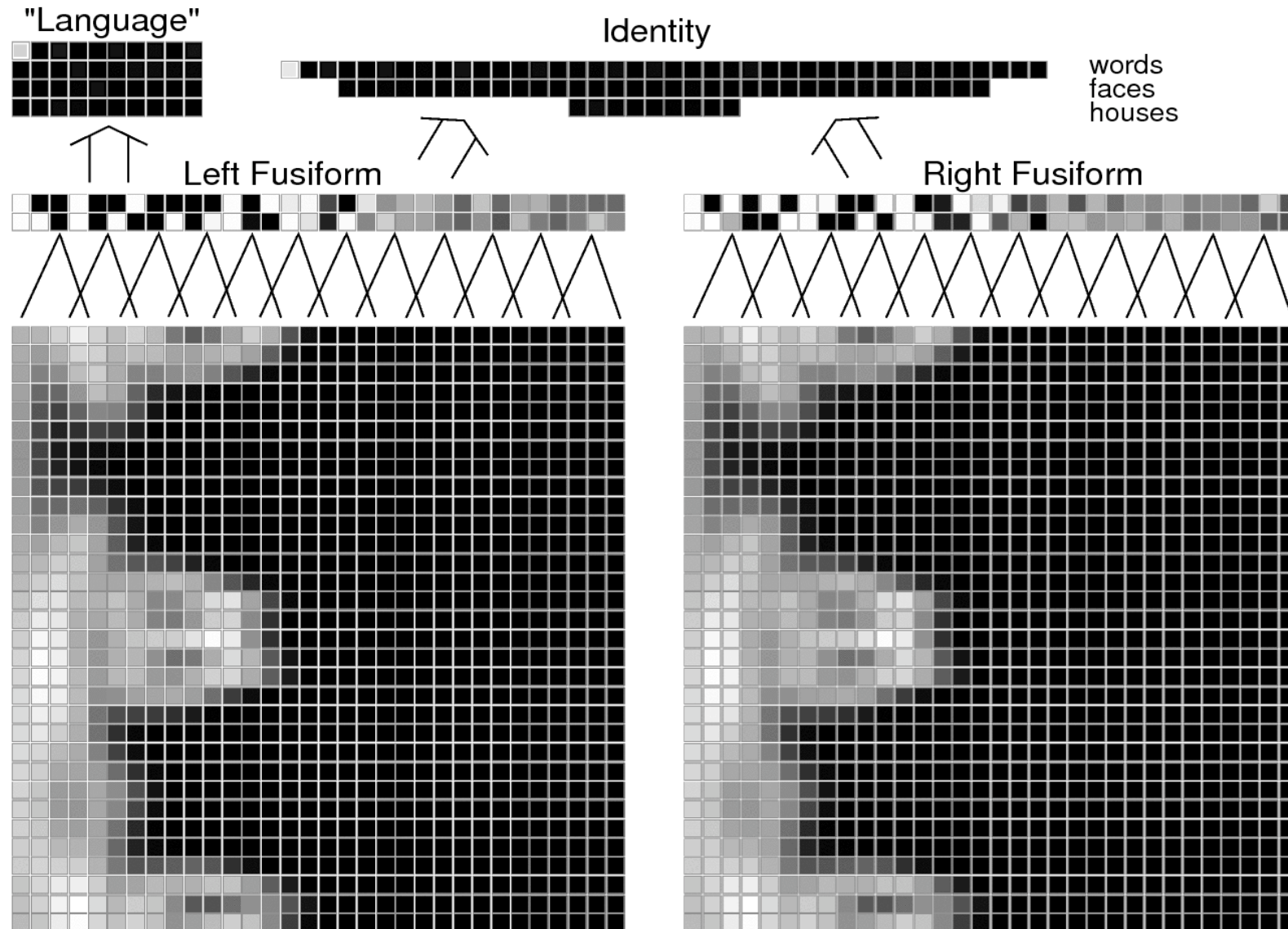
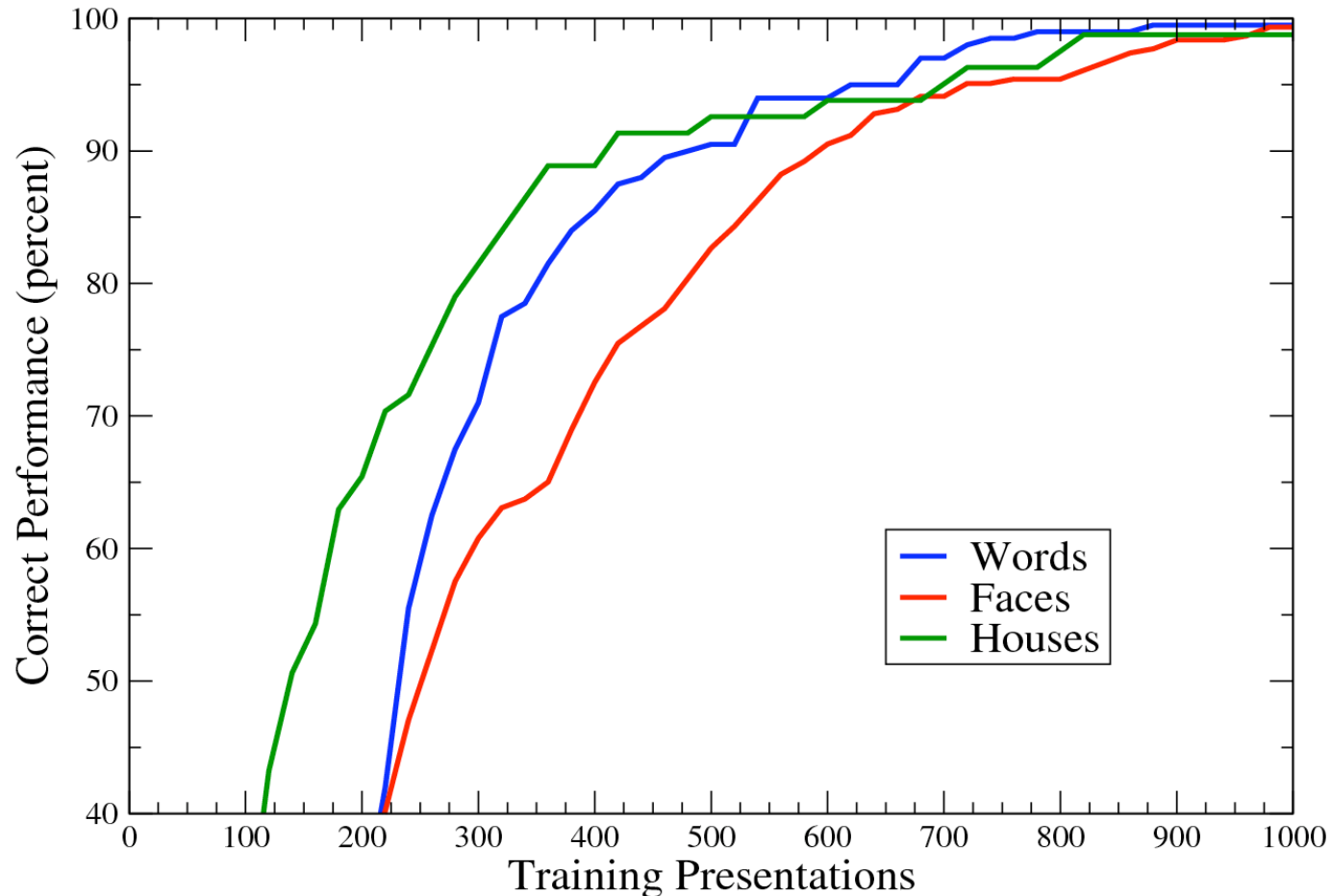
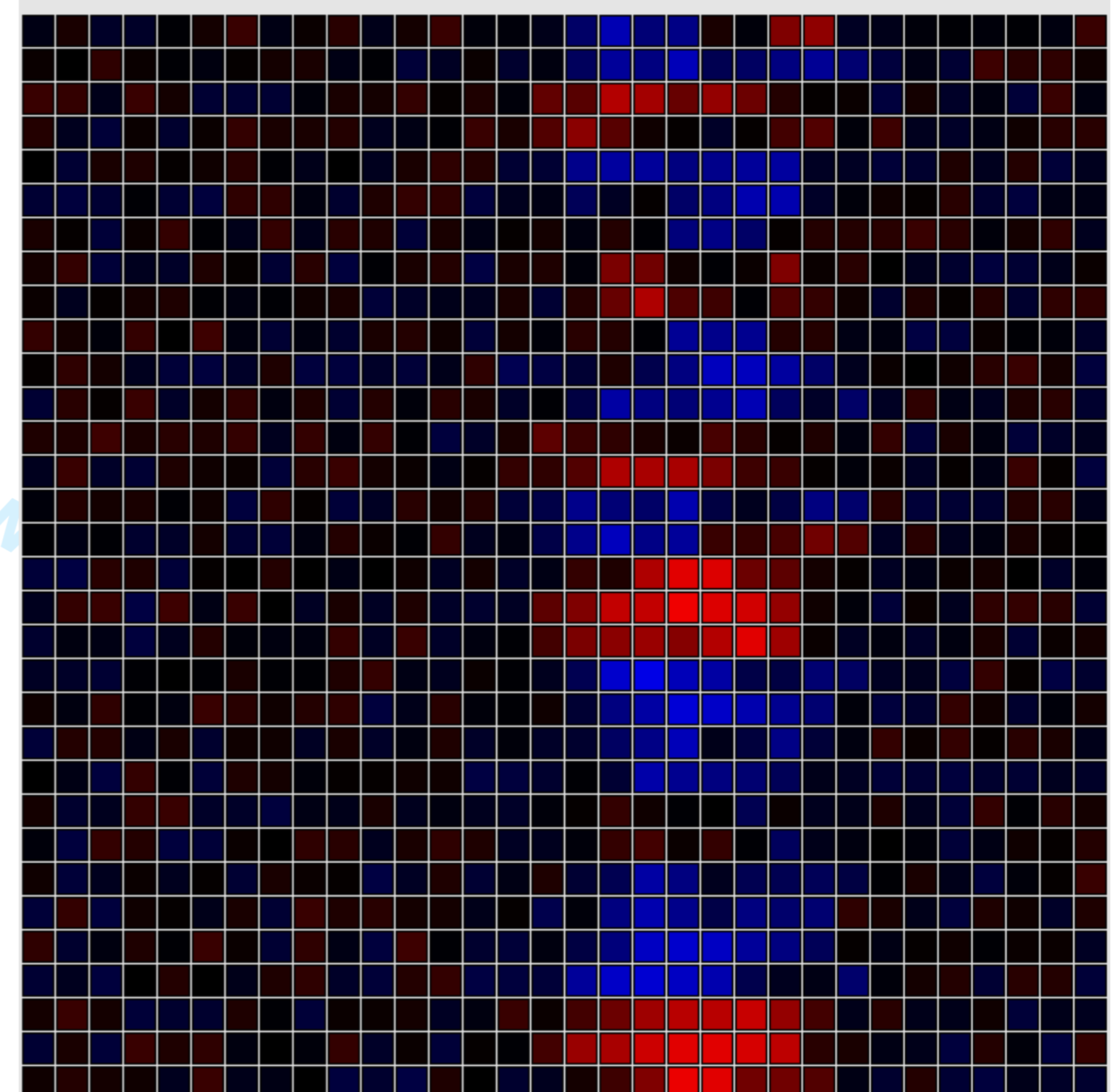
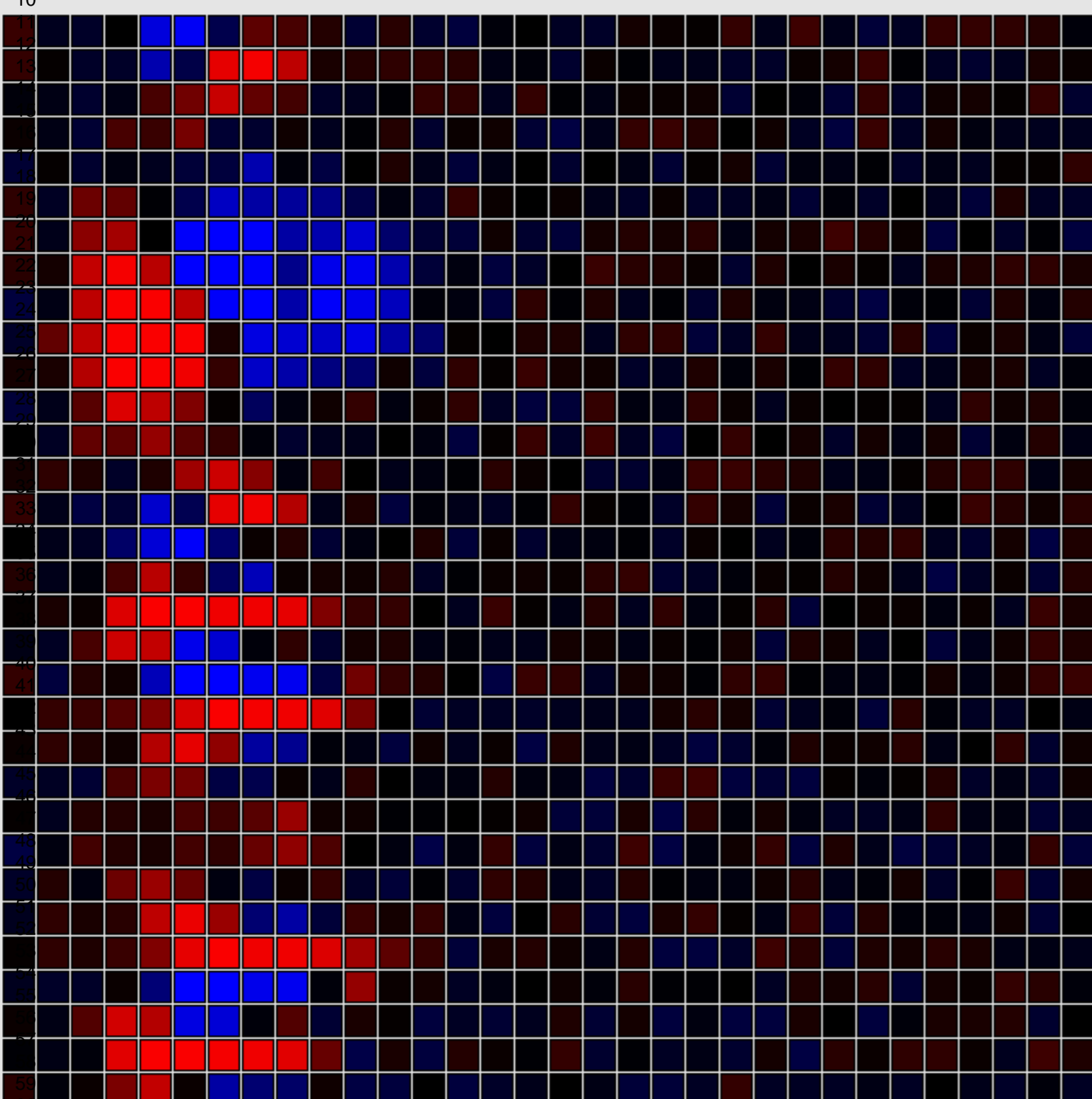
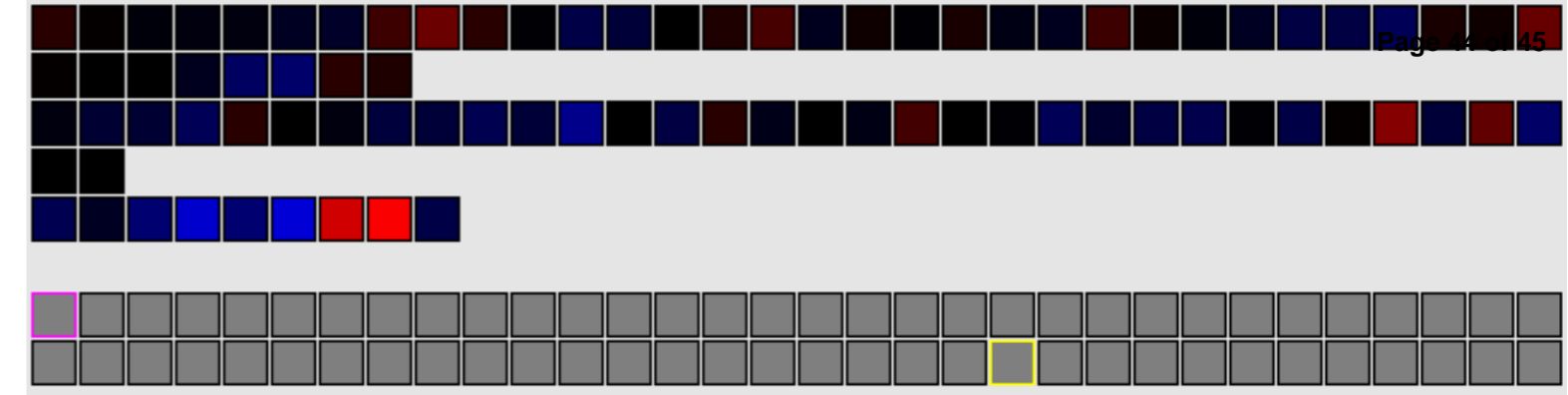


Figure 4





1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Figure 6

