

Annual Review of Vision Science

Reading: The Confluence
of Vision and Language

Jason D. Yeatman^{1,2,3} and Alex L. White^{1,2,4}

¹Graduate School of Education, Stanford University, Stanford, California 93405, USA;
email: jyeatman@stanford.edu

²Division of Developmental-Behavioral Pediatrics, Stanford University School of Medicine,
Stanford, California 94305, USA

³Department of Psychology, Stanford University, Stanford, California 94305, USA

⁴Department of Neuroscience and Behavior, Barnard College, New York, New York 10027, USA

**ANNUAL
REVIEWS CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annu. Rev. Vis. Sci. 2021. 7:487–517

First published as a Review in Advance on
June 24, 2021

The *Annual Review of Vision Science* is online at
vision.annualreviews.org

<https://doi.org/10.1146/annurev-vision-093019-113509>

Copyright © 2021 by Annual Reviews.
All rights reserved

Keywords

reading, visual word form area, VWFA, dyslexia, word recognition, retinotopy

Abstract

The scientific study of reading has a rich history that spans disciplines from vision science to linguistics, psychology, cognitive neuroscience, neurology, and education. The study of reading can elucidate important general mechanisms in spatial vision, attentional control, object recognition, and perceptual learning, as well as the principles of plasticity and cortical topography. However, literacy also prompts the development of specific neural circuits to process a unique and artificial stimulus. In this review, we describe the sequence of operations that transforms visual features into language, how the key neural circuits are sculpted by experience during development, and what goes awry in children for whom learning to read is a struggle.

What an astonishing thing a book is. It's a flat object made from a tree with flexible parts on which are imprinted lots of funny dark squiggles. But one glance at it and you're inside the mind of another person, maybe somebody dead for thousands of years. Across the millennia, an author is speaking clearly and silently inside your head, directly to you.

—Carl Sagan, *Cosmos*

1. INTRODUCTION

The goal of written language is to represent spoken language with a visual code. The challenge for the human visual system is to transform the complex array of features in a written word into sound and meaning. The literate brain accomplishes this challenge within a couple of hundred milliseconds (Bentin et al. 1999, Joo et al. 2021, Nobre et al. 1994, Salmelin et al. 2000, Sereno & Rayner 2003). In regions of the lateral temporal and frontal cortex implicated in language processing (referred to below as the language cortex), neural responses evoked by a written word closely approximate those evoked by a spoken word (Pugh et al. 2013, Rueckl et al. 2015). The process of word recognition begins in visual cortex but ultimately depends on rapid communication between cortical regions specialized for processing visual, auditory, and linguistic information.

While evolution endowed the human brain with specialized circuits for spoken language, written language is a relatively recent invention of human societies. Thus, literacy must emerge from circuits that evolved for other purposes, demonstrating the power of experience-dependent plasticity to sculpt neural circuits for unique functions. Children learn to read through thousands of hours of instruction and practice; the neural circuitry of reading develops in response to education. This makes reading unique with respect to many other functions of the visual system. For example, compare visual word recognition and face recognition. In skilled readers, the two abilities have many similarities. Deriving meaning from words on a page of text can be as effortless as recognizing a friend's face in a crowd. Moreover, in both cases, rapid recognition depends on specialized regions that selectively process each category of visual stimuli (Dehaene & Cohen 2011, Grill-Spector & Weiner 2014, Kanwisher 2010, Wandell et al. 2012, Yeatman et al. 2013).

However, there are important differences between the two skills. The ability to rapidly recognize faces is important for many mammalian species. Face recognition arguably depends on dedicated circuits that have been conserved across evolution (Kanwisher 2010, Kanwisher et al. 1997, McKone et al. 2012), although this idea is not without debate (Arcaro et al. 2017, 2019; Livingstone et al. 2017; Tarr & Cheng 2003; Tarr & Gauthier 2000). Face-selective regions in the ventral visual stream exist very early in life (Deen et al. 2017), and while their development certainly depends on experience (Arcaro et al. 2017), it does not require explicit instruction and training. In contrast, word recognition is uniquely human, and visual word-selective regions develop as a consequence of instruction. Children will not achieve literacy without systematic training and practice. Thus, the study of literacy provides a natural experiment to understand how experience and learning interact with processing constraints that are naturally present in the human visual system.

Much of this review focuses on a word-selective region of ventral occipitotemporal cortex (VOTC) that forms the gateway between the visual and language systems: the visual word form area (VWFA). The notion of a cortical module specialized for representing the visual form of words has a rich and contentious history (Bub et al. 1993). Jules Déjerine (1891) described the case of a patient who suffered a stroke and completely lost the ability to read. Déjerine chose the term word blindness, or alexia, to describe the patient's condition because spatial vision was intact; the patient was able to recognize common objects and navigate the world but was unable to read even a single word. Déjerine invoked the idea of a visual word form center that developed with literacy

Visual word form area (VWFA): a region (or collection of subregions) in VOTC that selectively responds to written words and mediates word recognition

to store information about orthography. That assertion sparked fierce resistance from the eminent neurologist Carl Wernicke. Wernicke described Déjerine's model as indefensible and grossly schematic (see Bub et al. 1993, p. 539). Thus began a debate about whether learned representations of visual word forms are at the foundation of literacy and reside in a particular brain region that connects vision and language (Bub et al. 1993, Eggert 1977, Wernicke 1906). The debate has continued into the twenty-first century (Dehaene & Cohen 2011; Price & Devlin 2003, 2011).

Early neuroimaging studies sought to localize cortical regions that encode visual word forms. In the early 2000s, it became clear that a swath of cortex extending from the posterior occipitotemporal sulcus (OTS) to roughly the midpoint of the fusiform gyrus displayed many of the properties of Déjerine's visual word form center. The term VWFA was coined to describe this region that selectively responds to visual words (Cohen et al. 2000, 2002). Recent studies have found that the VWFA is actually composed of at least two subregions (VWFA-1 and VWFA-2) that each perform distinct computations on words (Lerma-Usabiaga et al. 2018, White et al. 2019c). Modern techniques have allowed for a much more detailed understanding of the coding principles in the VWFA and its subregions than Déjerine could have anticipated. However, Déjerine's hypothesis has stood the test of time. Thus, we continue to use the term VWFA to refer to the entire word-selective portion of VOTC, but we also draw distinctions between the functional and anatomical properties of the VWFA subregions (VWFA-1/2) where appropriate.

Throughout this review, we often implicate a hierarchy of processing stages, implemented in distinct brain regions, that transform visual input into language. However, while it is tempting to view reading as an assembly line, in which the visual system performs a series of operations that are sent down the line to the language system, reciprocal connectivity between visual and language areas plays a key role in word recognition (Seidenberg & McClelland 1989). Even recognizing an isolated word depends on interactions between vision and language. For example, recent evidence suggests that frontal lobe language regions are activated in synchrony with, or potentially even earlier than, VOTC (Cornelissen et al. 2009, Wheat et al. 2013, Woodhead et al. 2014). The VWFA receives signals from early visual cortex; is sensitive to image properties that are ubiquitous across visual cortex, such as image contrast (Kay & Yeatman 2017); and is modulated by selective spatial attention to particular locations in the visual field (White et al. 2019c). However, the VWFA is also sensitive to linguistic properties of written words, such as the sounds represented by the letters (Glezer et al. 2016, 2019; McCandliss et al. 2003). Moreover, top-down signals activate the VWFA during many speech perception tasks (Dehaene et al. 2011, Pugh et al. 2013). Thus, the VWFA is the intersection of vision and language. The anatomy, physiology, and development of this region of VOTC is at the foundation of the uniquely human capacity for literacy.

The neural circuitry that transforms vision into language can only be understood in combination with models of reading behavior. In Section 2, we review key observations of reading behavior and the sensory and cognitive limitations on word recognition. We then return to a detailed characterization of the neural circuitry that underlies this behavior.

2. OVERVIEW OF READING BEHAVIOR

In spoken language, words are delivered serially, via changes in air pressure over time. Even the phonemes that constitute each word are delivered sequentially. By contrast, in written language, many words and their component letters are delivered all at once, via changes in light intensity across a surface. Nonetheless, it is clear that one cannot comprehend a passage of text all at once. There are many reasons for this, but the first major bottleneck is imposed by the constraints of peripheral vision. To illustrate this phenomenon, fixate your gaze on the red dot on the line of text in **Figure 1a**. You will find that there is a small window in the central visual field within which

Orthography:
the conventions that determine how a particular language is written, including the letters, statistics of letter combinations, and rules of spelling

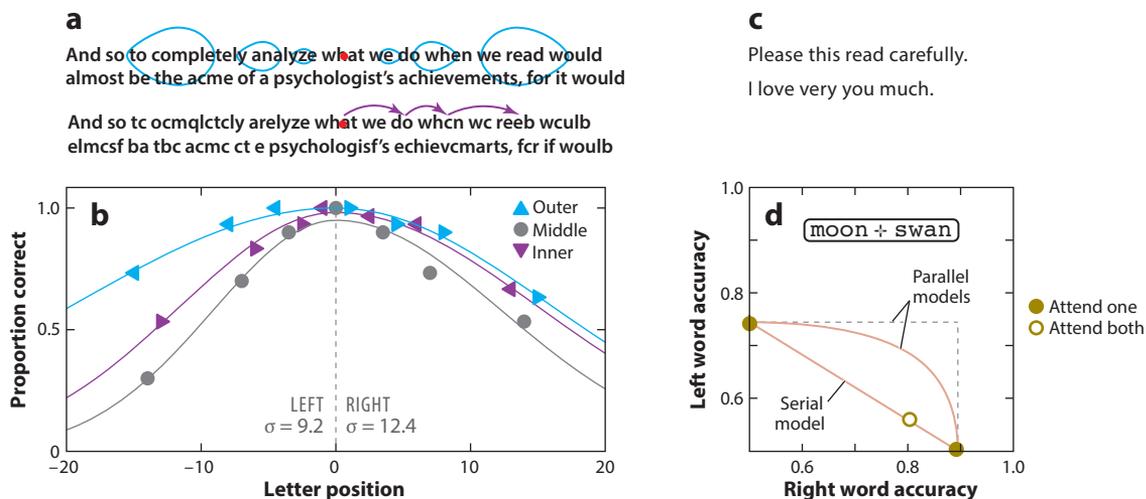


Figure 1

Crowding, the visual span, and serial versus parallel models of word recognition. (a) A quote from Huey (1908, p. 6). The red dot represents the current point of fixation. Crowding zones (light blue shapes) are wider than they are tall (radial or tangential anisotropy; see Toet & Levi 1992) and extend farther away from the fovea than toward it to account for the greater interference caused by flankers that are more eccentric than the target (Bouma 1973). In the next two lines of text, individual letters that have flankers within their crowding zones have been randomly replaced with other letters that are indistinguishable when crowded (silent substitution; see Pelli et al. 2007). The two passages should appear very similar when fixating the red dot in either one. The purple arrows indicate the saccades that a reader might make. (b) Visual span measurements in one subject (author A.L.W.) using a modified version of the trigram task from Legge et al. (2001). Trigrams (three-letter combinations) were presented for 117 ms along the horizontal meridian. This plot is comparable to Legge et al. (2001, figure 5), showing superior accuracy for the outer letter and strong crowding of the middle letter. A split Gaussian fit to the accuracy measures gives the width of the visual span and reveals the classic hemifield asymmetry for letter recognition. The left and right sigma parameters printed on the graph are for the middle letter. (c) Two sentences. Reread them until you understand the argument for parallel processing. Panel adapted from Snell & Grainger (2019b). (d) Average attention operating characteristic showing accuracy of semantic categorization of nouns presented to the left and right of fixation. Accuracy matches the serial model prediction ($N = 15$). Panel adapted with permission from White et al. (2019c).

letters and words are recognizable. Outside of that window, letters appear blurry and jumbled. Nonetheless, skilled readers learn to work within the constraints of their visual system to sample text at a nearly optimal rate for comprehension: 200–400 words per minute (Rayner et al. 2016).

This section provides an overview of how information arranged on a page gets funneled through the visual system and into the language system. We first focus on spatial factors: how orthographic information is sampled across the visual field and how the eyes and attention move across text. We then focus on temporal factors: whether letters and words are processed serially or in parallel.

Experimental psychologists have studied how readers extract information from text since at least the 1870s. Based on subjective experience, many assumed that the eyes smoothly scan over a page. However, one of the first discoveries in the science of reading was that the eyes make many short jerky movements (given the French name saccades in 1879; see Javal 1990) down a line of text (Hering 1879, Lamare 1892, Wade & Tatler 2008). That discovery immediately raised an important question: How much text is processed during a period of fixation? Initial measures by Hering (1879) and Lamare (1892) suggested that the amount was quite small, with an average of 10 letters. In a later experiment, Huey (1900) flashed entire lines of text (cut from a psychology journal) for 16 ms. The average distance between the leftmost and rightmost letters that participants could read was approximately 11 letter spaces but varied greatly across trials and

participants (Huey 1908). It became apparent that the amount of text processed at each moment is much smaller than many supposed, in part because readers are not aware of their own eye movements. The above quote from Carl Sagan, therefore, is not quite accurate: It takes more than one glance at a page to get inside the mind of the writer.

2.1. Crowding as a Fundamental Bottleneck in Word Recognition

Many authors, from the eighteenth century to the present day (e.g., Rayner et al. 2016), supposed that reading proceeds in a rapid sequence of fixations because of poor acuity outside of the fovea. Acuity refers to the smallest details that can be resolved in a single shape. Acuity drops off quickly with eccentricity, being limited by optics, retinal anatomy, and cortical processing (Wandell 1995). If a letter is too small, then it is difficult to recognize due to poor acuity.

However, for recognizing words, acuity is not typically the limiting factor. In his book, Huey (1908) reports his own findings [and others by Erdmann & Dodge (1898)] that the initial and final letters of a long word are more legible than the interior ones, despite being at greater eccentricity where acuity is worse. At the time, researchers could not explain this observation.

The discovery of visual crowding offered an explanation (for reviews, see Levi 2008, Manassi & Whitney 2018, Pelli & Tillman 2008, Strasburger 2020). Bouma (1970) showed that the ability to recognize a single letter decreases as retinal eccentricity increases but also that, at each position, recognition is impaired by the presence of flanking letters to either side. The amount of interference depends on both (*a*) the target letter's eccentricity and (*b*) the distance between the target and flankers. Specifically, Bouma postulated that interference occurs if the distance between the letters is approximately less than half the eccentricity, that is, if the flankers lie within the crowding zone, as illustrated in **Figure 1a**. Moreover, the effect of a flanker is stronger if it lies to the outside (at greater eccentricity) of the target than if it lies to the inside of it (Mackworth 1965). In measuring these phenomena, Bouma (1973) confirmed and explained what Huey (1908) originally described: that the outer letters of a string are often recognized better because they are less crowded than the interior ones (see also Bernard & Castet 2019, Nazir et al. 1991).

Therefore, as we try to read words at greater distances from the current point of fixation, letters crowd each other more. Crowding impairs recognition before acuity does. This point was eloquently stated by Pelli & Tillman (2008), who proposed that visual recognition is only successful in an uncrowded window that is determined not by the size of elements but by the spacing between them.

The region of the visual field within which letters can be recognized during fixation has been termed the visual span (O'Regan et al. 1983). Legge and colleagues have standardized a method of measuring the visual span by having subjects report the identities of three letters presented side by side at varying distances from fixation (e.g., Legge et al. 2001). **Figure 1b** shows an example data set. The width of the span allows approximately 10 letters to be recognized accurately (>80% correct) under typical reading conditions, consistent with the nineteenth-century measurements reviewed above. Several studies have argued that the visual span is a front-end bottleneck that directly constrains reading rate because the span and reading rate are similarly affected by physical manipulations of the text (Legge et al. 2007, Yu et al. 2007). Pelli et al. (2007) also argued that the size of the visual span is explained entirely by crowding (contingent upon some bending of Bouma's rule).

The formulation of the visual span as an uncrowded window predicts that individuals with a wider span should read faster and make longer saccades [as suggested by Ehlers (1936)]. Indeed, the total area of the visual span predicts individual reading speed and increases from childhood into adulthood before decreasing in later age (Kwon et al. 2007, Liu et al. 2017). Two other

Crowding: the impairment to object recognition caused by visual elements that are too close; if the distance between shapes is below an eccentricity-dependent limit, then they cannot be recognized

Visual span: the spatial window within which letters in a horizontal line can be recognized without moving the eyes; the illegibility of letters outside of the span is largely explained by crowding

Hemifield

asymmetry: the difference in letter and word recognition ability between the right and left visual hemifields; most people, regardless of language, perform better in the right visual field

studies investigated whether the leftward and rightward extents of the visual span predict natural reading behavior, with mixed results (Frömer et al. 2015, Risse 2014).

The formulation of the visual span also predicts that word recognition in the periphery can be matched to the fovea by increasing letter size and spacing above the eccentricity-dependent acuity limit and crowding zone, respectively. However, this is not the case. In general, reading speed increases with letter size and spacing up to a critical print size and then plateaus at a maximum speed. The maximum speed in the periphery never matches what is possible at the fovea (Chung et al. 1998, Latham & Whitaker 1996, Legge et al. 2001). This is an unfortunate constraint for individuals who have lost foveal vision and must read with a preferred retinal locus in the periphery. However, there is some evidence that training can speed peripheral reading (He et al. 2013, Yu et al. 2018).

The relative inefficiency of word recognition in peripheral vision—despite scaling size and spacing—is still unexplained (Levi 2008, Pelli et al. 2007). It may be a signature of unique characteristics of the reading circuitry, such as the field of view available to the VWFA (Le et al. 2017) (**Figure 2b**). More generally, we might ask whether the constraints on reading are general properties of the visual system or specific consequences of training the system to transform visual shapes into language.

2.2. The Right Hemifield Advantage: A Unique Characteristic of Visual Word Recognition

One piece of evidence that the visual system develops specific mechanisms that process text is an asymmetry in reading performance between the right and left visual fields. Huey (1900) reported that, when a line of text is flashed briefly, and participants fixate a point in the middle of it, they often report more words to the right of fixation than to the left. Mishkin & Forgays (1952) found that single English words are better recognized in the right visual field than in the left. One explanation of these two observations proposes different modes of processing in the two cerebral hemispheres (Bub & Lewine 1988), with more efficient parallel processing of letter strings in the left hemisphere (Ellis 2004; but see Whitney & Lavidor 2004).

The hemifield asymmetry also applies to crowded letter recognition (e.g., Bouma 1973, Nazir et al. 1991) such that the visual span extends farther to the right than to the left of fixation (Legge et al. 2001), as shown in **Figure 1b**. This may explain the optimal viewing position: Word recognition is facilitated when the point of fixation is to the left of the word's center, so that more letters are to the right (Brysaert et al. 1996, Nazir et al. 1991, O'Regan & Jacobs 1992). Similarly, during natural reading, people tend to fixate to the left of word centers (Rayner 1998).

Critically, the right hemifield advantage is stimulus specific. Nonletter shapes and faces either show no asymmetry or the opposite asymmetry (Fontenot 1973, He et al. 2015, Heron 1957, Hines 1978, Leehey & Cahn 1979). Some have proposed that the right hemifield advantage for words is caused by a general advantage for high spatial frequencies in the left hemisphere (Sergent 1982), but that notion has been challenged (Ossowski & Behrmann 2015, Tádros et al. 2013).

Is the hemifield asymmetry in English due to experience reading left to right? Apparently, this is not the case: Several right-to-left scripts (Arabic, Hebrew, Urdu) also have a right hemifield superiority, although perhaps of a smaller magnitude (for reviews, see Almabruk et al. 2011, Siéroff & Haehnel-Benoliel 2015). Note that the most recent evidence contradicts the first two studies on the topic (Mishkin & Forgays 1952, Orbach 1952). Therefore, the field is moving toward the conclusion that the typical right hemifield superiority for reading is likely due to the typical left hemisphere dominance for language. Confirming this theory, individuals with anomalous language lateralization have anomalous asymmetries for word recognition (Hunter et al. 2007; Van der Haegen et al. 2011, 2013).

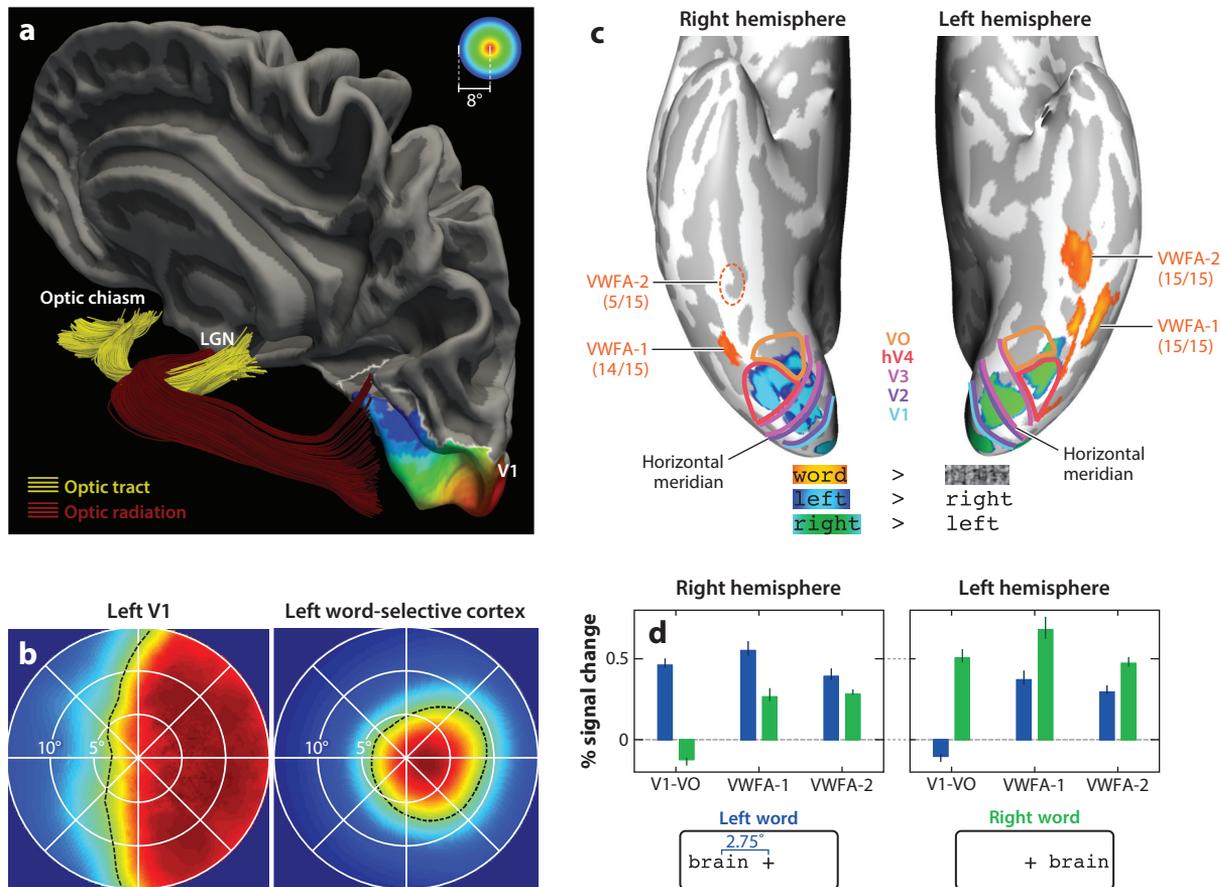


Figure 2

The front end of the reading circuitry. (a) The optic tract (yellow) and optic radiation (red) carry signals from the retina to the lateral geniculate nucleus (LGN) and then the primary visual cortex (V1). White matter tracts are taken from the Human Connectome Project (HCP) White Matter Atlas (Yeh et al. 2018). V1 is colored with eccentricity preferences from the HCP group average ($N = 181$) 7T retinotopy data set (Benson et al. 2018) mapped to the average cortical surface (Dale et al. 1999). (b) The field of view of left V1 and left word-selective cortex (Le et al. 2017). Images represent the portion of the visual field in which words evoke a response. The color at each visual field position represents the max of population receptive field (pRF) values across all voxels in the region. In V1, there is full coverage of the contralateral visual field. In word-selective cortex, coverage extends to approximately 10° in the right (contralateral) and approximately 5° in the left (ipsilateral) visual field. (c) Ventral view of the inflated cortical surface for one example subject from White et al. (2019c). A word presented at 2.75° left or right of fixation evokes a response at the corresponding location in each retinotopic map: V1, V2, V3, hV4, and VO. Blue shows responses within retinotopic cortex that are larger for words on the left than for those on the right, and green shows responses that are larger for words on the right than for those on the left. The orange color map represents responses that are larger for words than for phase-scrambled words. Two patches [visual word form area (VWFA)-1 and VWFA-2] were identifiable in the left hemisphere for all 15 subjects. Only one patch (VWFA-1) is usually identifiable in the right hemisphere. (d) Mean blood oxygen-level dependent (BOLD) responses ($N = 15$) to single words presented at $\pm 2.75^\circ$ (White et al. 2019c). Error bars = ± 1 standard error. Retinotopic cortex (averaged across V1-VO) responds positively only to contralateral words. VWFA-1 and -2 respond to words on both sides of fixation, with a contralateral preference.

For a thorough discussion of how stimulus-specific asymmetries arise, the reader is referred to Behrmann & Plaut (2020). An intriguing hypothesis is that recurrent connections with language centers help train the visual areas in the language-dominant hemisphere, which best represents the contralateral visual field, to process text (McCandliss et al. 2003). This hypothesis is in line

with data showing that the VWFA is colateralized with language (Gerrits et al. 2019, Haegen et al. 2012) and typically has enhanced representation of the right visual field (Le et al. 2017, White et al. 2019c) (**Figure 2b**).

2.3. Are Letters in a Word Processed Serially or in Parallel?

In this section, we discuss temporal properties of visual word recognition. One important theme in this area of research is the distinction between parallel and serial processes. It is clear that one cannot read all of the words in a paragraph in parallel, so there must be a serial mechanism to process a paragraph in chunks. How big are these chunks? Are they single letters, single words, or combinations of words?

The shapes of many letters are encoded in parallel by the retina and early visual cortex. Does that mean that all of the letters in a word are identified in parallel (Grainger et al. 2016) or in a rapid serial sweep (Whitney 2001)? Do individual letters even need to be identified, or can the word's form be recognized holistically? This is an old question, and attempts to answer it began with the word superiority effect.

Cattell (1886) observed that letters are more quickly identified when embedded in real words than when embedded in nonword letter strings or even when presented alone. The evidence for this word superiority effect was strengthened with an unbiased psychophysical task devised by Reicher (1969) and Wheeler (1970). The effect is difficult (but not impossible; see Whitney 2008) to explain if one assumes that letters are recognized one at a time, before whole words are recognized. One of the most influential models of word recognition, the interactive activation model (McClelland & Rumelhart 1981), proposes that all letters in a word are identified in parallel, activating representations of words that contain those letters; recurrent connections then enhance representations of the letters. Thus, knowledge of the whole word facilitates knowledge of the letters that it contains. The interaction between layers of the interactive activation model is in line with findings from neuroimaging studies of interactions between visual cortex and language cortex during the first 200 ms of word recognition (Carreiras et al. 2014). The role of word-level knowledge affecting letter perception is also demonstrated by the illusory letters phenomenon: When words are printed at a small size just at the edge of legibility, and the inner letters are replaced by nonletter shapes, real letters forming whole words are nonetheless seen (Jordan et al. 1999).

One interpretation of the word superiority effect is that words are recognized as wholes, without first identifying each letter independently. However, Pelli et al. (2003) demonstrated that the contrast energy required to identify a familiar word increases linearly with its length. This means that a word is not efficiently recognized as one pattern; it cannot be identified unless its component letters are identified.

Are all of the letters in a word actually identified in parallel? This is a difficult question to answer definitively. One line of research measures accuracy for single letters in briefly presented strings. The first letter is often recognized best, which some researchers believe is evidence of a serial sweep from left to right (Scaltritti & Balota 2013), but others disagree (Adelman et al. 2010, Tydgat & Grainger 2009). Another approach is to measure the time needed to recognize single words as a function of their length. A simple serial model would predict a linear increase in response time. However, for skilled readers, the number of letters in a word has little impact on recognition speed or accuracy as long as contrast is high, the word fits within the visual span (Legge et al. 2001), and the word is not in the left visual field (Ellis 2004). Note, however, that gaze durations during natural reading do increase with word length (Kliegl et al. 2004). Interestingly, lesions to the left VWFA cause a drastic increase in the effect of word length on recognition times

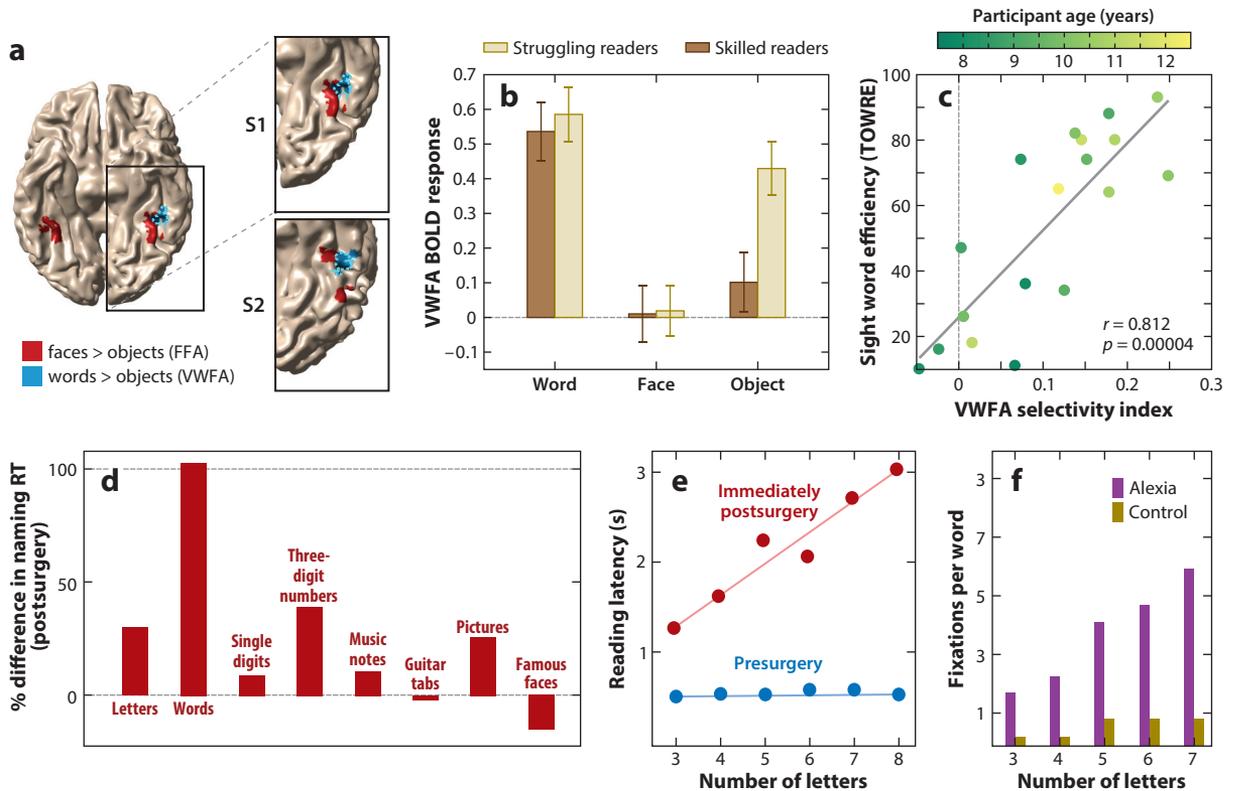


Figure 3

The VWFA is linked to reading behavior in children and adults. (*a–c*) Selectivity for words compared to objects in the VWFA is related to reading abilities in elementary school children (Kubota et al. 2019). (*a*) VWFA and FFA localizers in two representative children with strong reading abilities. The same region is identified when words are contrasted with faces or objects. (*b*) The BOLD response to words, faces, and objects in the VWFA for skilled readers and struggling readers. (*c*) As reading abilities increase, so does VWFA selectivity. VWFA selectivity is computed by comparing the response to words and objects. After surgical resection of the VWFA, (*d*) patients show selective deficits in naming words (Hirshorn et al. 2016), and (*e*) reading latency becomes linearly dependent on word length (Gaillard et al. 2006). (*f*) Patients with damage to VOTC from a stroke engage in letter-by-letter reading, making saccades between each letter in a word (Pflugshaupt et al. 2009). Abbreviations: BOLD, blood oxygen level–dependent; FFA, fusiform face area; RT, response time; VOTC, ventral occipitotemporal cortex; VWFA, visual word form area.

(**Figure 3e**), suggesting that the mature VWFA processes letter combinations in parallel (Gaillard et al. 2006, Pflugshaupt et al. 2009).

In summary, letter recognition is a necessary step toward word recognition. Notwithstanding some ambiguous data, many (but not all) models assume that letters in a word are processed in parallel, and the left VWFA is likely to be critical for doing so (Gaillard et al. 2006, Glezer et al. 2009, Hirshorn et al. 2016, Pflugshaupt et al. 2009, Strother et al. 2015).

2.4. Are Words on a Line Recognized Serially or in Parallel?

As mentioned above, readers make many saccades across a line of text, moving forward an average of seven letters each time. Between saccades, readers fixate for 250 ms on average. Both saccade lengths and fixation durations are highly variable, and many studies explore this variability to understand the cognitive processes at work (for reviews, see Radach & Kennedy 2013, Rayner

Perceptual span: the spatial window within which readers process visual information as they read; alterations to the text within the perceptual span affect reading behavior, while alterations outside of the span do not

1998, Rayner et al. 2016). Some words are fixated more than once, but approximately 30% of words are skipped (especially short function words). Skipping is one clue that, while the eye is jumping down the line, attention leads the way. This result is not specific to reading; an automatic attention shift to the peripheral target of an impending saccade occurs in many contexts (Rolfs & Carrasco 2012).

Experimenters have measured the portion of text that is attended with gaze-contingent reading paradigms that alter the text as the eyes move. For instance, in the moving window paradigm (McConkie & Rayner 1975), the text within a window of variable size around the current point of fixation is intact, but letters outside of that window are replaced by a mask. If the window is too narrow, then reading is disrupted. However, if the window is the same size as—or larger than—the amount of text being processed by the reader, then reading proceeds as normal. The smallest window that does not affect reading is the perceptual span, a measure that is quite different from the visual span (Frey & Bosse 2018). The perceptual span in English extends only 3–4 characters to the left of fixation and 14–15 characters to the right (Rayner 1998). In Chinese, it is much narrower but also asymmetric to the right (Inhoff & Liu 1998). In Hebrew (Pollatsek et al. 1981) and Arabic (Jordan et al. 2014), that asymmetry is reversed (unlike the hemifield asymmetry for isolated word recognition). The perceptual span is believed to be influenced by the shifting of attention, planning of eye movements, and contextual factors in sentence comprehension, unlike the narrower and less asymmetric visual span.

A variety of other gaze-contingent manipulations have demonstrated that participants do begin processing words to the right of fixation before the eyes move (Schotter et al. 2012). The type of information that is extracted from these words (visual, orthographic, semantic, etc.) is a topic of current debate (Snell & Grainger 2019b). The central issue is whether multiple words are processed to a high level in parallel, or whether linguistic processing of the next word only begins after the current word has been recognized.

To account for a cornucopia of eye movement data, research teams around the world have built computational models that either assume serial processing of words via sequential attention shifts (Reichle et al. 2006) or allow parallel processing of words under flexibly diffuse attentional gradients (Engbert et al. 2005, Reilly & Radach 2006, Snell et al. 2018b). Each type of model can account for a wide range of eye movement phenomena, so discriminating between them has been difficult.

Psychophysical measures of word recognition can be of assistance. One line of research asks if multiple words presented simultaneously interact, facilitating or inhibiting judgments (Mullin & Egeth 1989, Shepherdson & Miller 2014). For instance, judgments of a briefly (170 ms) flashed target word are influenced by irrelevant flanking words (Snell et al. 2017, 2018a). When the words around the target form a sentence, the target is reproduced more accurately than when the sentence is scrambled (a sentence superiority effect) (Snell & Grainger 2017, Wen et al. 2019). One interpretation of this result is that the words were all processed simultaneously. Furthermore, as argued by Snell & Grainger (2019b), a serial model predicts that words in a sentence should be perceived in order, and misordered words should be easily detected. This is not always the case, as you may experience when following the instructions in **Figure 1c** (Mirault et al. 2018). In these paradigms, however, it is difficult to distinguish perceptual interactions from post-perceptual decision processes (e.g., quickly correcting syntax errors to get the gist of the sentence, using prior expectations), especially when the time allowed for perceptual processing is not strictly controlled.

One may then directly ask whether skilled readers are capable of recognizing two words at exactly the same time. White et al. (2018, 2019c, 2020) presented participants with two words on either side of fixation, flashed briefly and post-masked. On each trial, the participant reported the

semantic category of one or both words (in a two-alternative forced-choice task). On different trials, they were precued to attend to just the left word, just the right one, or both. The time between the words and masks was adjusted so that each participant could achieve approximately 80% correct with focused attention. The threshold onset asynchrony in such tasks has been found to be surprisingly brief, ranging from 32 to 134 ms (depending on stimulus eccentricity and mask type).

The question asked was whether, in the same amount of time, participants could categorize both words. A series of experiments showed that the answer is no. As shown in the attention operating characteristic in **Figure 1d**, accuracy was far below the predictions of two different parallel models. Accuracy matched a serial model that assumes that only one word can be fully processed on each trial, and that no semantic information is acquired about the other. Note that **Figure 1d** also shows the hemifield asymmetry reviewed in Section 3.2: Far greater accuracy is observed for words to the right of fixation (x axis) than for those to the left (y axis). The serial result is robust to variations in the display arrangement, masking stimuli, and task requirements but does not apply to judgments of the text color, which can be identified for two words in parallel (White et al. 2020).

Not all researchers agree as to the significance of these psychophysical results for natural reading (Snell & Grainger 2019a, White et al. 2019b). Nonetheless, the data demonstrate a serial bottleneck in lexical access: Two independent words cannot be simultaneously recognized as words. Therefore, although reading begins with massive parallel processing of the retinal image, a small portion of that image—perhaps just one word—must be selected by attention and funneled into the language system. The next section reviews the latest research on the brain circuits that implement that linkage between the visual and language systems.

3. OVERVIEW OF THE READING CIRCUITRY

3.1. The Front End of the Reading Circuitry

The front end of the reading circuitry is no different than that used for other visual functions. Light reflected (or emitted) by a page (or screen) of text is encoded as neural impulses in the retina. Signals from retinal ganglion cells are carried to the lateral geniculate nucleus of the thalamus by the optic nerve and optic tract, and then to the primary visual cortex (V1) by the optic radiations (**Figure 2a**).

Retinotopy is a fundamental organizing principle for much of the mammalian visual system. In V1, individual neurons are tuned to oriented edges at specific locations in the visual field and arranged in a map of visual space such that adjacent points on the cortical surface respond to stimuli at adjacent points in the visual field. As signals progress through the hierarchy of retinotopically organized visual regions (i.e., V2, V3, V4), individual neurons become sensitive to increasingly complex patterns and integrate over larger areas of the visual field (Gallant et al. 1993, Nandy et al. 2013, Pasupathy et al. 2020).

Computational models of human visual cortex derived from functional magnetic resonance imaging (fMRI) have elucidated similar organizational principles as those derived from single-unit physiology. For example, the population receptive field (pRF) model describes the blood oxygen level–dependent (BOLD) signal at the spatial scale of fMRI measurements (voxels approximately 0.5 to 3 mm on each side, containing tens to hundreds of thousands of neurons) as the summation of signals over a portion of the visual field (Dumoulin & Wandell 2008). A voxel's pRF is generally defined in terms of its center location (x and y coordinates) and size (σ) (Dumoulin & Wandell 2008, Wandell & Winawer 2015). The average pRF size increases from V1 to hV4, as does the complexity of the model required to predict responses (Kay et al. 2013a,b).

Retinotopy: the mapping of retinal input to neurons in the brain, which mostly preserves the spatial layout of the retinal image on the cortical surface

Population receptive field (pRF) model: a computational model that summarizes the fMRI signal based on signals from a particular location in the visual field

Voxel: pixels within a three-dimensional imaging volume; each voxel summarizes MRI signals within a cubic volume of a couple of millimeters

Since the advent of fMRI in the early 1990s, dozens of retinotopic maps have been delineated in the human brain (Wandell & Winawer 2010), tessellating most of the occipital lobe and large swaths of the temporal (Amano et al. 2009, Huk et al. 2002), parietal (Silver et al. 2005), and even frontal lobes (Mackey et al. 2017, Silver & Kastner 2009). Each retinotopic region has a map of the contralateral visual hemifield (Arcaro et al. 2009, Brewer et al. 2005, Wandell et al. 2007, Winawer et al. 2010).

Paralleling the decrease in perceptual resolution in peripheral vision, the amount of cortical surface area devoted to each point in visual space decreases as eccentricity increases, and pRF sizes grow larger (Harvey & Dumoulin 2011). Thus, a word presented 3° to the right of fixation will produce activation at the corresponding locations of each retinotopic map in the left hemisphere (**Figure 2c**). A word at 6° will activate smaller patches nearby on the cortical surface. The notion that stimuli in the periphery are represented by fewer neurons with larger, more overlapping receptive fields may help explain why behavioral measurements of acuity (Duncan & Boynton 2003, Song et al. 2015) and crowding (He et al. 2019, Levi et al. 1985, Pelli 2008) differ across the visual field. Much of that variation in perceptual and cortical resolution is inherited from the distribution of retinal ganglion cells (Kwon & Liu 2019) but also differs among retinotopic regions (Harvey & Dumoulin 2011).

Therefore, we can draw some tentative links between properties of visual cortex and the reading behavior reviewed in the previous section. The retinotopic organization of early visual cortex allows for multiple letters to be processed in parallel, at least initially, by different populations of neurons within each retinotopic map. The size of the visual span and the difficulty of word recognition outside of the fovea could be explained in large part by the general properties of neuronal encoding across the visual field (e.g., larger receptive fields and more spatial integration in the periphery). Alternatively, these behavioral phenomena may relate to the fact that word-selective cortex only receives signals from a relatively small portion of the central visual field, much less than do V1 and other retinotopic regions (Le et al. 2017). Other stimulus-specific factors yet to be explained by properties of retinotopic cortex include the right hemifield advantage for text, the serial bottleneck for word recognition, and controversial claims that crowding mechanisms differ for letters (Castet et al. 2017, Grainger et al. 2010).

3.2. A Hierarchy of Processing Stages: From Features to Letters to Words

Letters are configurations of oriented and curved line segments. Most models of reading begin with the supposition that visual cortex first encodes those elementary visual features that, in combination, define individual letters (Balota et al. 2006, Dehaene et al. 2005, Grainger et al. 2008). Time-resolved measurements made with electroencephalography (EEG), magnetoencephalography (MEG), and electrocorticography (ECoG) demonstrate a temporal sequence beginning with sensitivity to basic visual features between 80 ms and 100 ms after stimulus onset (Marinkovic et al. 2003), progressing to sensitivity to letters between 120 ms and 160 ms (Thesen et al. 2012), and sensitivity to letter combinations and words at approximately 200 ms and later (Hirshorn et al. 2016). Spatially resolved measurements made with fMRI have revealed a sequence of regions in VOTC that selectively respond to letters and words (Cohen et al. 2000, 2002; Lerma-Usabiaga et al. 2018; Strother et al. 2015; Yeatman et al. 2013). Neurological studies have confirmed that this area of cortex, termed the VWFA for its specialized role in reading, is critical for rapid and automatic word recognition (Gaillard et al. 2006, Hirshorn et al. 2016, Pflugshaupt et al. 2009). Separate from the VWFA, posterior regions are selective for letters but without a preference for the letter combinations that constitute words (James et al. 2005, Strother et al. 2015, Vinckier et al. 2007, Wong et al. 2009).

3.3. Category Selectivity in Ventral Temporal Cortex

The VWFA lies in the OTS, surrounded by other visual regions that serve important functions for object recognition at a high level of the ventral visual stream (Felleman & Van Essen 1991). Up until very recently, the organization of high-level visual cortex was frequently described as inconsistent or bewildering (Weiner & Grill-Spector 2012). Neurons in the macaque inferior temporal (IT) cortex that selectively respond to faces, hands, and common objects were first discovered in the late 1960s (Gross et al. 1969, 1972). At the time, the existence of face cells seemed so improbable that other labs did not seek to replicate this observation for over a decade [in fact, Charlie Gross waited to publish the discovery for years (Squire 2009)]. Much later, human fMRI studies elucidated the elegant correspondence between anatomy and function in VOTC (Weiner & Grill-Spector 2012).

In 1997, Kanwisher and colleagues (Kanwisher 2010, Kanwisher et al. 1997) demonstrated that a face-selective region of the human fusiform gyrus could be localized in most individuals by contrasting the fMRI response to images of faces versus the response to images of various common objects (a localizer experiment). They called this region the face area, later referred to as the fusiform face area (FFA), and proposed that innately specified modules are devoted to the computations required for processing special stimulus classes such as faces. The existence of a face area was controversial, but later studies combining fMRI localizers with single-unit physiology in macaques confirmed that a vast majority of neurons in face areas are selective for faces (Tsao et al. 2006). Furthermore, combining fMRI localizers with cortical stimulation in human surgical patients confirmed that perturbing face-selective cortex selectively disrupts perception of faces (Parvizi et al. 2012). In the years immediately following the discovery of the FFA, adjacent regions with selectivity for places and scenes (Epstein & Kanwisher 1998), bodies (Downing et al. 2001), and words (i.e., the VWFA) (Cohen et al. 2000) were described and named. The VWFA is, of course, the most important of these areas for reading, and in the intervening 20 years, we have learned much more about its role in word recognition by combining fMRI, EEG, MEG, psychophysics, and direct intracranial recordings.

3.4. Inside the Visual Word Form Area: A Hierarchy of Subregions

By the early 2000s, it was clear that category selectivity is a fundamental organizing principle of high-level visual cortex, but it took another decade to work out a detailed model of the functional architecture. By collecting high-resolution fMRI data and visualizing face- and body-selective responses on the cortical surface of individual subjects, Weiner & Grill-Spector (2010) found that each category was represented in multiple, interdigitated patches, rather than in a single region. When visualized on anatomical slices in the brain volume, the layout of these repeating regions appeared haphazard. However, when visualized on the cortical surface, each subregion was consistently localized to the same anatomical landmark in each subject. For example, the FFA is actually composed of two distinct subregions, one of which is located in the posterior fusiform (termed FFA-1 or pFus-faces) and the second (FFA-2 or mFus-faces) in the anterior tip of the midfusiform sulcus (**Figure 4**). A body-selective region lies between these two face-selective regions (Weiner & Grill-Spector 2011). White and colleagues (2019c) demonstrated that the classic VWFA, like other areas, is also composed of distinct patches. VWFA-1 (or pFus-Words; Lerma-Usabiaga et al. 2018), located in the posterior OTS, is immediately lateral to the posterior face- and limb-selective regions (Grill-Spector & Weiner 2014). VWFA-2 is more anterior in the OTS (but also spread across the fusiform gyrus and inferior temporal gyrus), and its location is lateral and anterior to FFA-2.

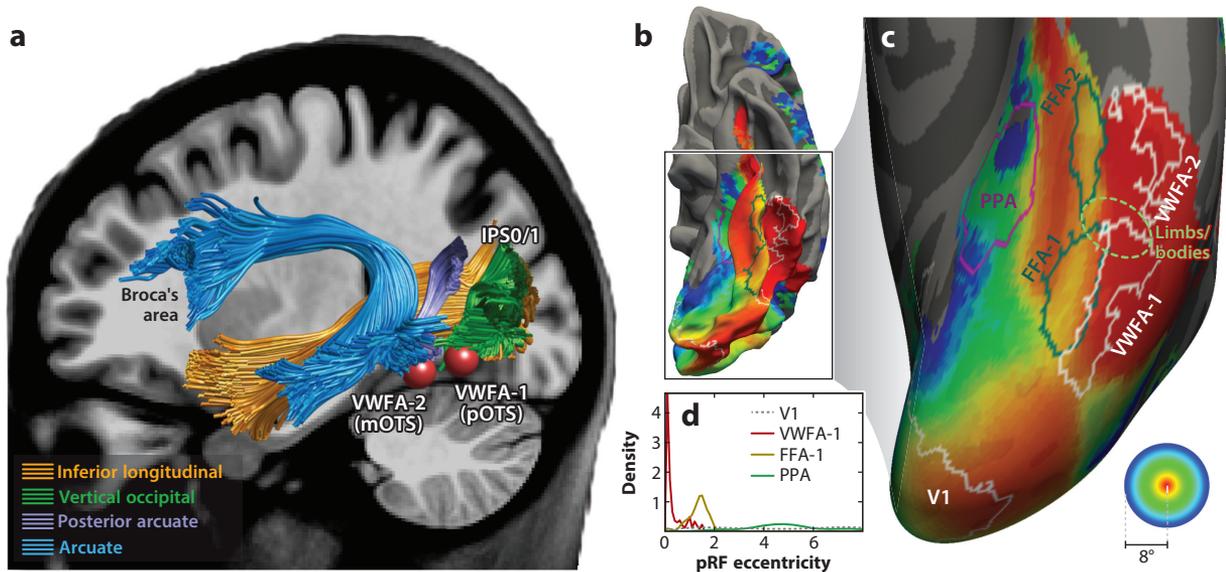


Figure 4

White matter pathways for reading. (a) VWFA-1 and VWFA-2 have different patterns of white matter connectivity. Activation peaks for words versus checkerboards in the posterior OTS (VWFA-1) and words versus pseudowords in the mid OTS (VWFA-2) are taken from Lerma-Usabiaga et al. (2018). The VOF (green) connects VWFA-1 to dorsal stream regions in the parietal lobe that are involved in visual attention (e.g., IPS 0/1 maps) (Takemura et al. 2016, Yeatman et al. 2014). The arcuate fasciculus (blue) connects VWFA-2 to language regions in lateral frontal cortex (e.g., Broca's area). The posterior segment of the arcuate (purple) connects VWFA-2 to regions in the superior temporal and inferior parietal lobe that are involved in the perception of speech sounds (Weiner et al. 2017b). (b) Eccentricity preferences based on the HCP group average retinotopy data set (Benson et al. 2018). The color map is thresholded to show regions where the pRF model $R^2 > 10\%$. An eccentricity gradient extends from the OTS, where pRFs are mostly centered within the central 1° (red), to the collateral sulcus, where pRFs are mostly centered $>6^\circ$ from fovea (green-blue). (c) Maximum probability maps of word-selective (VWFA), face-selective (FFA), and place-selective (PPA) regions (Grill-Spector & Weiner 2014, Weiner et al. 2018). The typical location of limb- and body-selective responses (also known as the fusiform body area; Peelen & Downing 2005) is between the face and word regions (Weiner & Grill-Spector 2011, Weiner et al. 2017a). (d) The distribution of pRF center eccentricities for V1 (which is relatively flat across the visual field), PPA (which has a bias towards the periphery), FFA-1 (which peaks around 2°), and VWFA-1 (which is heavily biased for the central 0.5°). Abbreviations: FFA, fusiform face area; HCP, Human Connectome Project; IPS, intra-parietal sulcus; mOTS, middle OTS; OTS, occipitotemporal sulcus; pOTS, posterior OTS; PPA, parahippocampal place area; pRF, population receptive field; V1, primary visual cortex; VOF, vertical occipital fasciculus; VWFA, visual word form area.

The notion of a processing hierarchy has long been a core principle in visual neuroscience (Felleman & Van Essen 1991) and computational models of vision (Riesenhuber & Poggio 1999). However, the order of operations across space and time is complex, with parallel and recurrent activity in multiple systems (Zeki 2016). The observation of multiple patches devoted to each stimulus category and arranged along the posterior–anterior axis of VOTC suggests the potential of hierarchical computations within each category or, at least, the division of computations into anatomically distinct regions that interact to derive meaning from a visual stimulus (Bao et al. 2020).

A recent ECoG study supports the notion of separable VOTC subregions involved in word recognition but challenges a strict hierarchical model that assumes a temporal order of activation from posterior to anterior regions. Direct neural recordings showed increasing selectivity for orthographic and lexical properties of words along the posterior–anterior axis, but selectivity for words over nonwords appeared in a relatively anterior midfusiform region before spreading to

more posterior regions (Woolnough et al 2021). It remains to be seen whether these subregions correspond exactly to the VWFA-1 and VWFA-2 described with fMRI.

Examinations of the white matter connections of VOTC also lend support to the hypothesis that VWFA subregions are anatomically distinct and fulfill distinct roles in the word-recognition process. The more posterior VWFA-1, which appears to encode visual properties of text (Kay & Yeatman 2017, Lerma-Usabiaga et al. 2018) and is modulated by spatial attention (White et al. 2019c), is located within the ventral terminations of the inferior longitudinal fasciculus (ILF) and vertical occipital fasciculus (VOF) (Kay & Yeatman 2017, Lerma-Usabiaga et al. 2018, Yeatman et al. 2013) (**Figure 4a**). The ILF carries signals from early visual cortex up the ventral stream. The VOF connects VOTC to V3A/B and intraparietal sulcus (IPS) 0/1 maps that are involved in control of spatial attention (Silver & Kastner 2009, Silver et al. 2005, Takemura et al. 2016, Yeatman et al. 2014). VWFA-2 is anterior to the VOF and is mostly within the ventral terminations of the arcuate fasciculus (Lerma-Usabiaga et al. 2018, Weiner et al. 2017b, Yeatman et al. 2014), although it also receives projections from the ILF (Grotheer et al. 2020). The arcuate fasciculus connects to Broca's area in lateral frontal cortex and has played a central role in theories of language for over a century (Catani & Mesulam 2008, Wernicke 1881). Thus, VWFA-2 is uniquely positioned to function as an intermediary between vision and language. In line with this perspective, VWFA-2 is uniquely sensitive to linguistic properties of text (Lerma-Usabiaga et al. 2018, White et al. 2019c).

4. THE REPRESENTATION OF VISUAL SPACE IN WORD-SELECTIVE CORTEX

4.1. Eccentricity Biases

Retinotopic maps have not been discovered in the OTS or fusiform gyrus, where patches of cortex are selective for words, faces, and other image categories (but see Kolster et al. 2010). Near these patches of cortex are retinotopic maps with very small foveal representations in scene-selective regions of the collateral sulcus and parahippocampal cortex (PHC-1 and PHC-2) (Arcaro et al. 2009, Lescroart & Gallant 2019).

The VWFA is not clearly retinotopic, but its lack of a topographic map of the visual field does not imply that its neurons do not have spatially tuned receptive fields. For example, Malach and colleagues (Hasson et al. 2002, Levy et al. 2001, Malach et al. 2002) discovered that lateral VOTC has a bias for foveal signals (within the central 1.5°), whereas medial VOTC has a bias for peripheral signals (greater than 6°). Thus, even though there is no topographic map of polar angle, there is a gradient of eccentricity preferences that spans centimeters of VOTC (**Figure 4b**). This finding implies that every VOTC region will have an eccentricity preference, which may determine its category preference. Under this hypothesis (Hasson et al. 2002), the VWFA is in the OTS, at the lateral boundary of VOTC, because that region has the strongest preference for foveal signals (**Figure 4b,c,d**).

This observation has stood the test of time. Face-selective regions are immediately medial to the VWFA (**Figure 4b,c**) and have a field of view that extends 2–3° further into the periphery than that of the VWFA (Kay et al. 2015, Malach et al. 2002, Silson et al. 2015) (**Figure 3d**). Face perception depends on foveal signals but also involves integration of information over a larger portion of the visual field; while a word typically spans 1–2° of visual angle, a face typically spans 4–8° in natural viewing (Grill-Spector et al. 2017). Some evidence even suggests that, as children become highly skilled readers, words compete with faces for the most foveal representations in VOTC (Gomez et al. 2018). Scene-selective regions are located in the collateral sulcus, the medial boundary of VOTC (Epstein & Kanwisher 1998, Epstein et al. 1999, Nasr et al. 2011). Scene-selective

Topographic map: an orderly projection of a stimulus dimension (e.g., eccentricity) onto the surface of the cortex; retinotopy is an example of a topographic map where the surface of the retina is projected onto the surface of the cortex

regions, in principle, must integrate information over large parts of the visual field, and indeed, they have the strongest bias toward information in the periphery (greater than 6° eccentricity) (Arcaro et al. 2009).

4.2. Spatial Invariance Versus Spatial Tuning

In this section, we discuss the question of the representation of visual space within word-selective cortex. Early reports concluded that the VWFA was invariant to spatial position because words presented in both the left and right hemifield evoked a strong response in the left-hemisphere OTS (Cohen et al. 2002). This was a surprising observation given how strictly selective other visual areas are for the contralateral visual field and, more generally, the ubiquity of retinotopy throughout visual cortex. Rauschecker and colleagues (2012) challenged the position invariance of the VWFA by presenting words at six locations along the horizontal and vertical meridians. They then trained a support vector machine to classify the stimulus location based on the pattern of responses evoked across all of the voxels in each visual region. Using the pattern of responses in the VWFA, they correctly classified the location of the word approximately 75% of the time. Although accuracy was higher in early visual cortex (>90% correct), these data clearly demonstrate that VWFA voxels systematically vary in terms of their sensitivity to words at specific locations in the visual field. Follow-up work demonstrated that the pRF model was a good fit to data from individual VWFA voxels and that, even though most voxels had pRF centers near fovea, there was substantial variability in pRF center and size across VWFA voxels (Le et al. 2017).

The discovery of position sensitivity within the VWFA demanded a revision to the prevailing hypothesis that the representation of word forms is invariant to position (Goebel 2012, Hannagan & Grainger 2013). What role does this spatial information play in the word-recognition process? One possibility is that information about retinal position is inherited from early visual cortex but does not play an essential role in VWFA computations (Hannagan & Grainger 2013). Alternatively, the VWFA might have more similarities to other visual areas than is commonly presumed; neurons within the VWFA might have receptive fields that only encode word forms at particular locations in the visual field (Goebel 2012, Rauschecker et al. 2012).

One way to test the functional significance of position information is through a manipulation of spatial attention. In early retinotopic areas, attending to a particular location in the visual field amplifies the response of neurons with receptive fields centered on the attended location (Brefczynski & DeYoe 1999, Gandhi et al. 1999, Pestilli et al. 2011, Silver et al. 2005).

White and colleagues (2019c) measured the effects of spatial attention in word-selective VOTC regions by adapting a classic spatial cueing paradigm to a word-recognition task. To determine each voxel's spatial tuning, they first presented single words 3° to the left or right of fixation. Consistent with the findings of Rauschecker and colleagues (2012), word-selective regions in both hemispheres responded to words on both sides of fixation, but with a clear preference for the contralateral visual field (**Figure 2c**). Moreover, in the posterior VWFA subregion (VWFA-1), voxels varied in their spatial preferences.

White and colleagues (2019c) then capitalized on each word-selective voxel's spatial tuning in an experiment manipulating spatial attention (using the behavioral paradigm reviewed above; see **Figure 1d**). Two words were presented simultaneously on each trial, one to the left and one to the right of fixation, and the subject was pre-cued to attend to the left side, the right side, or both. Voxels tuned to the left visual field responded more strongly when subjects attended to the left word, and voxels tuned to the right responded more strongly when subjects attended to the right word. In this respect, VWFA-1 voxels behave in a similar manner to retinotopic regions:

Each voxel has a preference for a particular location and is amplified when attention is allocated to that location. However, this experiment also found a striking and possibly unique feature of left-hemisphere VWFA-1. A spatial encoding model based on individual voxel activity suggested that the population of neurons in left VWFA-1 composes (at least) two spatial channels, one for words in the left visual field and one for words in the right. These two channels can be independently modulated by selective attention.

The notion of separable spatial channels in the posterior word-selective region also relates to the processing of a single word that is fixated centrally. The two halves of the word are initially processed in contralateral hemispheres of early visual cortex. Repetition suppression experiments have shown that a posterior word-selective region represents both halves separately, while an anterior region (lateralized to the left hemisphere) integrates information from both hemifields into a unified representation of a whole word (Strother et al. 2015, 2017). In this regard, these word-selective regions are quite unlike earlier retinotopic regions, where the visual hemifields are more strictly segregated into separate hemispheres.

4.3. From Parallel Processing of Visual Features to Serial Processing of Linguistic Features

White and colleagues (2019b,c) made another observation that offers a resolution to the controversy over whether the VWFA is invariant (Cohen et al. 2002) or sensitive (Le et al. 2017, Rauschecker et al. 2012) to the position of words in the visual field. The key was to analyze separate subregions of the VWFA. Much of the classic work on the VWFA relied on group averages, spatial smoothing, and visualizations within the volumetric coordinate space defined by a standard brain template. This group-average approach lacks anatomical precision, particularly in the case of regions that vary in size and position among subjects (Glezer & Riesenhuber 2013). Moreover, group averages mix the data across the boundaries of regions that are discernable in the brains of each individual (Wandell et al. 2012). In contrast, White and colleagues defined VWFA-1 and VWFA-2 in the brains of individual subjects. Only VWFA-1 (posterior OTS) showed clear spatial tuning that varied across voxels and supported the hypothesis of two spatial channels that could be modulated by selective attention. In VWFA-2, the pattern of activity was better fit by a model that assumes signals are integrated across the two hemifields into a single channel that processes only one attended word at a time.

Thus, Rauschecker et al. (2012) and White et al. (2019c) found unambiguous support for variation in spatial tuning across word-selective voxels in the posterior OTS of individual subjects (VWFA-1). In contrast, other studies that concluded that the VWFA is location invariant used group averages and data visualized within the brain volume. They were likely analyzing or mixing signals from the more anterior subregion (i.e., VWFA-2), where there is less position information. Future research must build on the fact that the classic VWFA is composed of (at least) two separable subregions, possibly organized in a hierarchy.

Given that (a) VWFA-1 has multiple spatial channels that encode words across the visual field, (b) VWFA-2 only has a single spatial channel, and (c) behavioral measures indicate that you can only recognize one word at a time, one might ask what occurs when a subject is cued to attend to two words simultaneously. Retinotopic areas encode multiple stimuli at different locations in parallel, and, when an individual is cued to attend to multiple locations in the visual field, attention amplifies the response at all of the attended locations, compared to when each location is ignored (Chen & Seidemann 2012, White et al. 2017; but see Pestilli et al. 2011). The same is true in VWFA-1: Attending to two words amplifies the response in voxels that are tuned to each location (White et al. 2019c). VWFA-2 is unaffected by the distribution of spatial attention (focus left, focus

right, or divide) in this context. The working model based on these data is that parallel signals from VWFA-1 converge into a bottleneck in VWFA-2, which only processes a single word at a time (White et al. 2019b). This model offers an explanation of the behavioral observation of a serial bottleneck in word recognition (White et al. 2020) and, more generally, an answer to the question of how lines of text are transformed into sequences of meaningful words.

5. WHAT DETERMINES THE LOCATION OF THE VISUAL WORD FORM AREA?

Next, we might ask why the VWFA is in this exact location. What is special about the left OTS for processing words? Given that words are a new cultural invention, and not an evolutionarily relevant category of visual stimuli, it is surprising that the VWFA always develops in the same anatomical location across languages, cultures, and writing systems.

There are two related but distinct hypotheses to explain the emergence of the VWFA in the OTS. The neuronal recycling hypothesis, proposed by Dehaene & Cohen (2007), posits that new cultural inventions find their neuronal niche in circuits that evolved for closely related functions. Specifically, as children learn to read, word recognition competes with face recognition for cortical territory. As literacy training reorganizes a patch of left VOTC to selectively process words, face-selective responses become increasingly right lateralized (Behrmann & Plaut 2020, Dehaene et al. 2011). Why are responses to words in VOTC typically left lateralized? Interestingly, there is a correlation between the lateralization of written word responses in VOTC and the lateralization of higher-level language processing (Broca's area) (Cai et al. 2010, Gerrits et al. 2019, Haegen et al. 2012). Individuals with right-lateralized language regions have stronger responses to written words in right VOTC. Therefore, in line with the neuronal recycling hypothesis, the VWFA may develop in regions with the capacity to process complex objects at the fovea and with direct connections to language centers, although the lateralized hemispheric dominance is not total (Behrmann & Plaut 2020).

A related but distinct hypothesis explains the topography of VOTC as an object space whereby different dimensions of object perception might specify the optimal location for different category-selective regions. Broadly encapsulated by the object space hypothesis, a variety of dimensions have been proposed, including fovea versus periphery (Hasson et al. 2002, Malach et al. 2002), animate versus inanimate, small versus large real world size (Konkle & Oliva 2012), high versus low spatial frequency, and various shape characteristics (Op de Beeck et al. 2019).

Regarding the neuronal recycling hypothesis, some recent studies have called into question the supposition that words compete with faces for cortical territory (see Hervais-Adelman et al. 2019, Huettig et al. 2018). Two studies that followed children longitudinally from preschool reported that word selectivity emerges from a region that is otherwise not selective for any specific category (Dehaene-Lambertz et al. 2018, Saygin et al. 2016). Moreover, a cross-sectional study of elementary school children suggested that word-selective cortex emerges within a region that would otherwise process a variety of different types of objects (Kubota et al. 2019) (**Figure 3a–c**). Thus, although it is clear that literacy changes the organization of VOTC, it is still debatable whether literacy competes with any one specific category for space in VOTC (but see Nordt et al. 2020). In fact, in comparisons of literate adults and illiterate adults, literacy is associated with larger-amplitude responses to other categories of visual stimuli beyond words (Hervais-Adelman et al. 2019).

Regarding the object space hypothesis, Bao and colleagues (2020) collected fMRI data and single-unit recordings in VOTC while rhesus macaques viewed 1,224 images of objects. They were able to explain the pattern of neuronal responses by placing each object in a two-dimensional

space: Along one dimension, objects ranged between spiky and stubby, and along the other, they ranged from animate to inanimate. Even in patches of VOTC not clearly selective for categories like faces, bodies, or scenes, neurons seemed to be sorting objects according to these two dimensions.

Based on rough homologies between macaques and humans, the object space model from Bao et al. (2020) predicts that the VWFA emerges in a portion of VOTC that is tuned for inanimate, stubby objects. Correspondingly, the FFA arises where there is tuning for round or stubby animate objects. This principle is in line with new discoveries of protomaps that are present at birth and predict the future locations of category-selective regions (Arcaro & Livingstone 2017, Arcaro et al. 2019, Livingstone et al. 2017). In fact, when young macaques are trained to recognize letters, they develop a letter-selective region in a cortical location that is selective for foveal signals from birth and roughly homologous to the VWFA in humans (Srihasam et al. 2014).

The concept of an object space (or protomap) provides an appealing and parsimonious description of VOTC, but there are certainly other important factors, such as the pattern of structural connections with language cortex, that also constrain the location of the VWFA. For example, in congenitally blind individuals, the OTS still develops a selectivity for written language in the form of braille (Bedny 2017, Kim et al. 2017, Reich et al. 2011). Moreover, the location of major white matter tracts measured in preschool children predicts where the future VWFA will emerge as these children learn to read (Saygin et al. 2016).

Thus, the location in visual cortex that is reorganized to support literacy is likely determined by several constraints, including (a) preferences for foveal signals, (b) predisposition for certain visual features, and (c) the presence of long-range white matter connections to requisite visual and language regions.

6. WHAT IS COMPUTED IN WORD-SELECTIVE CORTEX AND HOW DOES IT SUPPORT WORD RECOGNITION?

Central to the question in the title of this section is the century-old debate over the existence of a specialized visual representation of written language that holds knowledge about letters, their combinatorial statistics, and the identity of words (Bub et al. 1993, Wernicke 1906). Contrary to the hypothesis that word-selective cortex encodes specific properties of text that mediate word recognition, Wernicke (1906) and more recent critics (Price & Devlin 2003, 2011) have argued that the visual system only encodes generic properties of visual form. A recent ECoG study tackled this question by identifying word-selective electrodes located on the left mid-fusiform gyrus and then stimulating these electrodes while subjects named letters, words, faces, and various pictures. This study found that electrical stimulation applied to word-selective cortex specifically interfered with naming letters and words but had no effect on naming faces and other pictures (Hirshorn et al. 2016). Moreover, after surgical resection, patients displayed selective deficits for naming words (**Figure 3d**). Analyzing the response of word-selective electrodes revealed sensitivity to statistical regularities of letter combinations such that a pattern classifier could detect whether text had common versus uncommon letter combinations within 200 ms of stimulus onset. In a later time window, a pattern classifier was sensitive to individual word identity (Hirshorn et al. 2016).

These direct neural recordings agree with a growing body of fMRI studies suggesting that the VWFA (likely VWFA-2) encodes word identity (Glezer et al. 2009, 2015; Riesenhuber & Glezer 2017). Taken together, these studies suggest that learning new written words creates lexical entries that are stored in the VWFA and can be differentiated from letter strings that differ by even a

single letter (Glezer et al. 2009, Riesenhuber & Glezer 2017). The VWFA is therefore sensitive to the frequency statistics of letter combinations that we encounter during reading, as well as to the identity of words.

7. WHY DO SOME CHILDREN STRUGGLE TO LEARN TO READ?

Reading is a complex behavior that depends on coordinated activity across a collection of brain systems. Thus, one might hypothesize that a deficit in any component of the reading circuitry, or in the white matter connections that carry signals between disparate regions, would interfere with the process of learning to read. However, many of the prominent theories of reading disability, also known as developmental dyslexia, posit a single underlying mechanism that explains why so many children struggle to read (Goswami 2015).

The predominant view is that dyslexia is a language-based learning disability, specifically caused by deficits in phonological processing (Snowling 1998, Vellutino et al. 2004, Wagner & Torgesen 1987). Phonological processing refers to encoding, accessing, and manipulating the sounds of spoken language—a core set of skills that are undoubtedly important for learning to read (Boets et al. 2013, Wagner & Torgesen 1987). However, this view is not without controversy. Other research groups assert that deficits in visual–spatial attention (Franceschini et al. 2012, Vidyasagar & Pammer 2010) and/or the magnocellular visual pathway (Eden et al. 1996, Stein 2001) are the core deficits in dyslexia. The magnocellular theory of dyslexia emerged in the 1980s from observations that people with dyslexia perform poorly on psychophysical tasks that tap magnocellular function (Lovegrove et al. 1980), such as visual motion discrimination (Demb et al. 1998). Moreover, postmortem dissections revealed differences in the size and morphology of neurons in the magnocellular layers of the lateral geniculate nucleus in a few individuals with dyslexia (Livingstone et al. 1991).

Although debates over competing core deficit models of dyslexia persist, many researchers are moving to accept the multifactorial nature of dyslexia (O'Brien & Yeatman 2021, Pennington 2006). Like other developmental disorders (e.g., autism spectrum disorders), dyslexia is unlikely to have a single cause common to all children. Instead, dyslexia should be conceptualized as the outcome of a combination of different risk factors. Under this model, poor phonological processing (Wolf & Bowers 1999), general issues with spoken language (Catts et al. 2017), difficulties controlling spatial attention (White et al. 2019a), elevated visual crowding (Joo et al. 2018, Martelli et al. 2009), and deficiencies at various stages of visual processing (Joo et al. 2017, Talcott et al. 2002) all confer risk for reading difficulties. Under this multifactorial framework, dyslexia is the probabilistic outcome of a collection of deficits of varying severity (Catts et al. 2017, O'Brien & Yeatman 2021, Pennington et al. 2012).

Regarding the underlying physiology, many of the components of the reading circuitry reviewed in this article are implicated in dyslexia. In fMRI studies, left VOTC (including the VWFA) is the most commonly reported location of anomalous activity in people with dyslexia (Richlan et al. 2011, Wandell et al. 2012). In children with dyslexia, it is often impossible to identify the VWFA, likely due to the lack of a word-selective response (Kubota et al. 2019) (**Figure 3a–c**). Moreover, the developmental trajectories of the arcuate fasciculus and ILF differ in children with dyslexia, and this difference persists into adulthood (Boets et al. 2013, Wandell & Yeatman 2013, Yeatman et al. 2012). Recent data even suggest that VOTC, and the associated white matter connections, are different in preschool children prior to formal reading instruction (Centanni et al. 2019, Langer et al. 2015, Vandermosten et al. 2015). Fortunately, many of these differences in brain structure (Huber et al. 2018) and function (Barquero et al. 2014) show high levels of plasticity

when children receive intensive, evidence-based intervention programs. Understanding how development of the visual system relates to development of the rest of the reading circuitry will, we hope, pave the way for new, personalized interventions for those who struggle to learn to read.

8. CONCLUSION

As you read this review, your brain performed an incredible feat: It transformed a complex sequence of visual symbols into a narrative at 200–300 words per minute. You made thousands of rapid eye movements and shifts of attention to process the text in small chunks. That was only possible because, early in your life, education sculpted specialized circuits that lie at the intersection of vision and language. In that respect, reading is unique compared to other visual functions: Literacy requires systematic instruction and practice over a protracted period of development. However, much of the circuitry for reading is shared across a myriad of other visual functions and inherits both the capacities and constraints that are present at many stages of the visual system. Therefore, a nuanced understanding of reading is by definition an interdisciplinary endeavor, drawing on research spanning vision science, psychology, neuroscience, linguistics, and education. In this review, we cover the rich history of the scientific study of reading to provide context for the most recent discoveries linking the structure and function of the brain's reading circuitry to the behavior that it supports. With the compendium of methods that are now available to psychologists and neuroscientists, the field is poised to answer many of the outstanding questions that have intrigued reading researchers since the nineteenth century.

FUTURE ISSUES

1. Why is reading in the peripheral visual field so slow and difficult? If the problem were entirely due to generic visual constraints, increasing letter size and spacing should equate peripheral and foveal reading, but it does not.
2. How many distinct word-selective regions exist in high-level visual cortex, and what are their functions?
3. Should the subcomponents of the VWFA be conceptualized as distinct regions organized in a hierarchy or as a gradient with gradual changes in function across the cortical surface?
4. To what extent do processing constraints in the visual system contribute to reading difficulties in people with dyslexia, and what are the implications for intervention?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We would like to thank Brian Wandell, Kevin Weiner, and Garikoitz Lerma-Usabiaga for helpful comments and discussion on the article and materials for the figures. We would like to thank Noah Benson for assistance with the Human Connectome Project retinotopy data and associated figures. This work was funded by National Institutes of Health (NIH) National Institute of Child Health

and Human Development grants R01HD09586101 and R21HD092771 and a Jacobs Foundation Research Fellowship to J.D.Y. and NIH National Eye Institute grant K99 EY029366 to A.L.W.

LITERATURE CITED

- Adelman JS, Marquis SJ, Sabatos-DeVito MG. 2010. Letters in words are read simultaneously, not in left-to-right sequence. *Psychol. Sci.* 21(12):1799–801
- Almabruk AAA, Paterson KB, McGowan V, Jordan TR. 2011. Evaluating effects of divided hemispheric processing on word recognition in foveal and extrafoveal displays: the evidence from Arabic. *PLOS ONE* 6(4):e18131
- Amano K, Wandell BA, Dumoulin SO. 2009. Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. *J. Neurophysiol.* 102(5):2704–18
- Arcaro MJ, Livingstone MS. 2017. A hierarchical, retinotopic proto-organization of the primate visual system at birth. *eLife* 6:e26196
- Arcaro MJ, McMains SA, Singer BD, Kastner S. 2009. Retinotopic organization of human ventral visual cortex. *J. Neurosci.* 29(34):10638–52
- Arcaro MJ, Schade PF, Livingstone MS. 2019. Universal mechanisms and the development of the face network: What you see is what you get. *Annu. Rev. Vis. Sci.* 5:341–72
- Arcaro MJ, Schade PF, Vincent JL, Ponce CR, Livingstone MS. 2017. Seeing faces is necessary for face-domain formation. *Nat. Neurosci.* 20(10):1404–12
- Balota DA, Yap MJ, Cortese MJ. 2006. Visual word recognition: the journey from features to meaning (a travel update). In *Handbook of Psycholinguistics*, ed. M Traxler, M Gernsbacher, pp. 285–375. Amsterdam: Elsevier. 2nd ed.
- Bao P, She L, McGill M, Tsao DY. 2020. A map of object space in primate inferotemporal cortex. *Nature* 583:103–8
- Barquero LA, Davis N, Cutting LE. 2014. Neuroimaging of reading intervention: a systematic review and activation likelihood estimate meta-analysis. *PLOS ONE* 9(1):e83668
- Bedny M. 2017. Evidence from blindness for a cognitively pluripotent cortex. *Trends Cogn. Sci.* 21(9):637–48
- Behrmann M, Plaut DC. 2020. Hemispheric organization for visual object recognition: a theoretical account and empirical evidence. *Perception* 49(4):373–404
- Benson NC, Jamison KW, Arcaro MJ, Vu AT, Glasser MF, et al. 2018. The Human Connectome Project 7 Tesla retinotopy dataset: description and population receptive field analysis. *J. Vis.* 18(13):23
- Bentin S, Mouchetant-Rostaing Y, Giard MH, Echallier JF, Pernier J. 1999. ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *J. Cogn. Neurosci.* 11(3):235–60
- Bernard J-B, Castet E. 2019. The optimal use of non-optimal letter information in foveal and parafoveal word recognition. *Vis. Res.* 155:44–61
- Boets B, Op de Beeck HP, Vandermosten M, Scott SK, Gillebert CR, et al. 2013. Intact but less accessible phonetic representations in adults with dyslexia. *Science* 342(6163):1251–54
- Bouma H. 1970. Interaction effects in parafoveal letter recognition. *Nature* 226(5241):177–78
- Bouma H. 1973. Visual interference in the parafoveal recognition of initial and final letters of words. *Vis. Res.* 13(4):767–82
- Brefczynski JA, DeYoe EA. 1999. A physiological correlate of the “spotlight” of visual attention. *Nat. Neurosci.* 2(4):370–74
- Brewer AA, Liu J, Wade AR, Wandell BA. 2005. Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nat. Neurosci.* 8(8):1102–9
- Brysbaert M, Vitu F, Schroyens W. 1996. The right visual field advantage and the optimal viewing position effect: on the relation between foveal and parafoveal word recognition. *Neuropsychology* 10(3):385–95
- Bub DN, Arguin M, Lecours AR. 1993. Jules Dejerine and his interpretation of pure alexia. *Brain Lang.* 45(4):531–59
- Bub DN, Lewine J. 1988. Different modes of word recognition in the left and right visual fields. *Brain Lang.* 33(1):161–88

- Cai Q, Paulignan Y, Brysbaert M, Ibarrola D, Nazir TA. 2010. The left ventral occipito-temporal response to words depends on language lateralization but not on visual familiarity. *Cereb. Cortex* 20(5):1153–63
- Carreiras M, Armstrong BC, Perea M, Frost R. 2014. The what, when, where, and how of visual word recognition. *Trends Cogn. Sci.* 18(2):90–98
- Castet E, Descamps M, Denis-Noël A, Colé P. 2017. Letter and symbol identification: no evidence for letter-specific crowding mechanisms. *J. Vis.* 17(11):2
- Catani M, Mesulam M. 2008. The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex* 44(8):953–61
- Cattell JM. 1886. The time it takes to see and name objects. *Mind* 11(41):63–65
- Catts HW, McIlraith A, Bridges MS, Nielsen DC. 2017. Viewing a phonological deficit within a multifactorial model of dyslexia. *Read. Writ.* 30(3):613–29
- Centanni TM, Norton ES, Ozernov-Palchik O, Park A, Beach SD, et al. 2019. Disrupted left fusiform response to print in beginning kindergartners is associated with subsequent reading. *NeuroImage Clin.* 22:101715
- Chen Y, Seidemann E. 2012. Attentional modulations related to spatial gating but not to allocation of limited resources in primate V1. *Neuron* 74(3):557–66
- Chung ST, Mansfield JS, Legge GE. 1998. Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vis. Res.* 38(19):2949–62
- Cohen L, Dehaene S, Naccache L, Lehericy S, Dehaene-Lambertz G, et al. 2000. The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123(2):291–307
- Cohen L, Lehericy S, Chochon F, Lemer C, Rivaud S, Dehaene S. 2002. Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain* 125(Pt. 5):1054–69
- Cornelissen PL, Kringelbach ML, Ellis AW, Whitney C, Holliday IE, Hansen PC. 2009. Activation of the left inferior frontal gyrus in the first 200 ms of reading: evidence from magnetoencephalography (MEG). *PLOS ONE* 4(4):e5359
- Dale AM, Fischl B, Sereno MI. 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage* 9(2):179–94
- Deen B, Richardson H, Dilks DD, Takahashi A, Keil B, et al. 2017. Organization of high-level visual cortex in human infants. *Nat. Commun.* 8:13995
- Dehaene S, Cohen L. 2007. Cultural recycling of cortical maps. *Neuron* 56(2):384–98
- Dehaene S, Cohen L. 2011. The unique role of the visual word form area in reading. *Trends Cogn. Sci.* 15(6):254–62
- Dehaene S, Cohen L, Sigman M, Vinckier F. 2005. The neural code for written words: a proposal. *Trends Cogn. Sci.* 9(7):335–41
- Dehaene S, Pegado F, Braga LW, Ventura P, Nunes Filho G, et al. 2011. How learning to read changes the cortical networks for vision and language. *Science* 330(6009):1359–64
- Dehaene-Lambertz G, Monzalvo K, Dehaene S. 2018. The emergence of the visual word form: longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLOS Biol.* 16(3):e2004103
- Déjerine J. 1891. Sur un cas de cécité verbale avec agraphie suivi d'autopsie. *Mém. Soc. Biol.* 3:197–201
- Demb JB, Boynton GM, Best M, Heeger DJ. 1998. Psychophysical evidence for a magnocellular pathway deficit in dyslexia. *Vis. Res.* 38(11):1555–59
- Downing PE, Jiang Y, Shuman M, Kanwisher N. 2001. A cortical area selective for visual processing of the human body. *Science* 293(5539):2470–73
- Dumoulin SO, Wandell BA. 2008. Population receptive field estimates in human visual cortex. *NeuroImage* 39(2):647–60
- Duncan RO, Boynton GM. 2003. Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron* 38(4):659–71
- Eden GF, VanMeter JW, Rumsey JM, Maisog JM, Woods RP, Zeffiro TA. 1996. Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. *Nature* 382(6586):66–69
- Eggert GH. 1977. *Wernicke's Works on Aphasia: A Sourcebook and Review*. Berlin: de Gruyter
- Ehlers H. 1936. The movements of the eyes during reading. *Acta Ophthalmol.* 14:56–63
- Ellis A. 2004. Length, formats, neighbours, hemispheres, and the processing of words presented laterally or at fixation. *Brain Lang.* 88(3):355–66

- Engbert R, Nuthmann A, Richter EM, Kliegl R. 2005. SWIFT: a dynamical model of saccade generation during reading. *Psychol. Rev.* 112(4):777–813
- Epstein R, Harris A, Stanley D, Kanwisher N. 1999. The parahippocampal place area: recognition, navigation, or encoding? *Neuron* 23(1):115–25
- Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. *Nature* 392:598–601
- Erdmann B, Dodge R. 1898. *Psychologische Untersuchungen über das Lesen, auf Experimenteller Grundlage*. Halle, Ger.: Max Niemeyer
- Felleman DJ, Van Essen DC. 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1(1):1–47
- Fontenot DJ. 1973. Visual field differences in the recognition of verbal and nonverbal stimuli in man. *J. Comp. Physiol. Psychol.* 85(3):564–69
- Franceschini S, Gori S, Ruffino M, Pedrolli K, Facoetti A. 2012. A causal link between visual spatial attention and reading acquisition. *Curr. Biol.* 22(9):814–19
- Frey A, Bosse M-L. 2018. Perceptual span, visual span, and visual attention span: three potential ways to quantify limits on visual processing during reading. *Vis. Cogn.* 26(6):412–29
- Frömer R, Dimigen O, Niefind F, Krause N, Kliegl R, Sommer W. 2015. Are individual differences in reading speed related to extrafoveal visual acuity and crowding? *PLOS ONE* 10(3):e0121986
- Gaillard R, Naccache L, Pinel P, Clemenceau S, Volle E, et al. 2006. Direct intracranial, fMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron* 50(2):191–204
- Gallant JL, Braun J, Van Essen DC. 1993. Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* 259(5091):100–3
- Gandhi SP, Heeger DJ, Boynton GM. 1999. Spatial attention affects brain activity in human primary visual cortex. *PNAS* 96(6):3314–19
- Gerrits R, Van der Haegen L, Brysbaert M, Vingerhoets G. 2019. Laterality for recognizing written words and faces in the fusiform gyrus covaries with language dominance. *Cortex* 117:196–204
- Glezer LS, Eden G, Jiang X, Luetje M, Napoliello E, et al. 2016. Uncovering phonological and orthographic selectivity across the reading network using fMRI-RA. *NeuroImage* 138:248–56
- Glezer LS, Jiang X, Luetje MM, Napoliello EM, Kim J, et al. 2019. An fMRI-adaptation study of phonological and orthographic selectivity to written words in adults with poor reading skills. *Brain Lang.* 191:1–8
- Glezer LS, Jiang X, Riesenhuber M. 2009. Evidence for highly selective neuronal tuning to whole words in the “visual word form area.” *Neuron* 62(2):199–204
- Glezer LS, Kim J, Rule J, Jiang X, Riesenhuber M. 2015. Adding words to the brain’s visual dictionary: novel word learning selectively sharpens orthographic representations in the VWFA. *J. Neurosci.* 35(12):4965–72
- Glezer LS, Riesenhuber M. 2013. Individual variability in location impacts orthographic selectivity in the “visual word form area.” *J. Neurosci.* 33(27):11221–26
- Goebel R. 2012. Position coding in the visual word form area. *PNAS* 109(24):9226–27
- Gomez J, Natu V, Jeska B, Barnett M, Grill-Spector K. 2018. Development differentially sculpts receptive fields across early and high-level human visual cortex. *Nat. Commun.* 9(1):788
- Goswami U. 2015. Sensory theories of developmental dyslexia: three challenges for research. *Nat. Rev. Neurosci.* 16(1):43–54
- Grainger J, Dufau S, Ziegler JC. 2016. A vision of reading. *Trends Cogn. Sci.* 20(3):171–79
- Grainger J, Rey A, Dufau S. 2008. Letter perception: from pixels to pandemonium. *Trends Cogn. Sci.* 12(10):381–87
- Grainger J, Tydgate I, Isselé J. 2010. Crowding affects letters and symbols differently. *J. Exp. Psychol. Hum. Percept. Perform.* 36(3):673–88
- Grill-Spector K, Weiner KS. 2014. The functional architecture of the ventral temporal cortex and its role in categorization. *Nat. Rev. Neurosci.* 15(8):536–48
- Grill-Spector K, Weiner KS, Kay K, Gomez J. 2017. The functional neuroanatomy of human face perception. *Annu. Rev. Vis. Sci.* 3:167–96
- Gross CG, Bender DB, Rocha-Miranda CE. 1969. Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science* 166(3910):1303–6

- Gross CG, Rocha-Miranda CE, Bender DB. 1972. Visual properties of neurons in inferotemporal cortex of the macaque. *J. Neurophysiol.* 35(1):96–111
- Grotheer M, Yeatman J, Grill-Spector K. 2020. White matter fascicles and cortical microstructure predict reading-related responses in human ventral temporal cortex. *NeuroImage* 227:117669
- Hannagan T, Grainger J. 2013. The lazy visual word form area: computational insights into location-sensitivity. *PLOS Comput. Biol.* 9(10):e1003250
- Harvey BM, Dumoulin SO. 2011. The relationship between cortical magnification factor and population receptive field size in human visual cortex: constancies in cortical architecture. *J. Neurosci.* 31(38):13604–12
- Hasson U, Levy I, Behrmann M, Hendler T, Malach R. 2002. Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34(3):479–90
- He D, Wang Y, Fang F. 2019. The critical role of V2 population receptive fields in visual orientation crowding. *Curr. Biol.* 29(13):2229–36.e3
- He Y, Legge GE, Yu D. 2013. Sensory and cognitive influences on the training-related improvement of reading speed in peripheral vision. *J. Vis.* 13(7):14
- He Y, Scholz JM, Gage R, Kallie CS, Liu T, Legge GE. 2015. Comparing the visual spans for faces and letters. *J. Vis.* 15(8):7
- Hering E. 1879. *Über Muskelgerausche des Auges*. Vienna: Gerold
- Heron W. 1957. Perception as a function of retinal locus and attention. *Am. J. Psychol.* 70(1):38–48
- Hervais-Adelman A, Kumar U, Mishra RK, Tripathi VN, Guleria A, et al. 2019. Learning to read recycles visual cortical networks without destruction. *Sci. Adv.* 5(9):eaax0262
- Hines D. 1978. Visual information processing in the left and right hemispheres. *Neuropsychologia* 16(5):593–600
- Hirshorn EA, Li Y, Ward MJ, Richardson RM, Fiez JA, Ghuman AS. 2016. Decoding and disrupting left midfusiform gyrus activity during word reading. *PNAS* 113(29):8162–67
- Huber E, Donnelly PM, Rokem A, Yeatman JD. 2018. Rapid and widespread white matter plasticity during an intensive reading intervention. *Nat. Commun.* 9(1):2260
- Huetting F, Kolinsky R, Lachmann T. 2018. The culturally co-opted brain: how literacy affects the human mind. *Null* 33(3):275–77
- Huey EB. 1900. On the psychology and physiology of reading. I. *Am. J. Psychol.* 11(3):283–302
- Huey EB. 1908. *The Psychology and Pedagogy of Reading*. New York: Macmillan
- Huk AC, Dougherty RF, Heeger DJ. 2002. Retinotopy and functional subdivision of human areas MT and MST. *J. Neurosci.* 22(16):7195–205
- Hunter ZR, Brysbaert M, Knecht S. 2007. Foveal word reading requires interhemispheric communication. *J. Cogn. Neurosci.* 19(8):1373–87
- Inhoff AW, Liu W. 1998. The perceptual span and oculomotor activity during the reading of Chinese sentences. *J. Exp. Psychol. Hum. Percept. Perform.* 24(1):20–34
- James KH, James TW, Jobard G, Wong ACN, Gauthier I. 2005. Letter processing in the visual system: different activation patterns for single letters and strings. *Cogn. Affect. Behav. Neurosci.* 5(4):452–66
- Javal E. 1990. Essay on the physiology of reading. *Ophthalmic Physiol. Opt.* 10(4):381–84
- Joo SJ, Donnelly PM, Yeatman JD. 2017. The causal relationship between dyslexia and motion perception reconsidered. *Sci. Rep.* 7(1):4185
- Joo SJ, Tavabi K, Caffarra S, Yeatman JD. 2021. Automaticity in the reading circuitry. *Brain Lang.* 215:104906
- Joo SJ, White AL, Strodtman DJ, Yeatman JD. 2018. Optimizing text for an individual's visual system: the contribution of visual crowding to reading difficulties. *Cortex* 103:291–301
- Jordan TR, Almabruk AAA, Gadalla EA, McGowan VA, White SJ, et al. 2014. Reading direction and the central perceptual span: evidence from Arabic and English. *Psychon. Bull. Rev.* 21(2):505–11
- Jordan TR, Thomas SM, Scott-Brown KC. 1999. The illusory-letters phenomenon: an illustration of graphemic restoration in visual word recognition. *Perception* 28:1413–16
- Kanwisher N. 2010. Functional specificity in the human brain: a window into the functional architecture of the mind. *PNAS* 107(25):11163–70
- Kanwisher N, McDermott J, Chun MM. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17(11):4302–11
- Kay KN, Weiner KS, Grill-Spector K. 2015. Attention reduces spatial uncertainty in human ventral temporal cortex. *Curr. Biol.* 25(5):595–600

- Kay KN, Winawer J, Mezer A, Wandell BA. 2013a. Compressive spatial summation in human visual cortex. *J. Neurophysiol.* 110(2):481–94
- Kay KN, Winawer J, Rokem A, Mezer A, Wandell BA. 2013b. A two-stage cascade model of BOLD responses in human visual cortex. *PLOS Comput. Biol.* 9(5):e1003079
- Kay KN, Yeatman JD. 2017. Bottom-up and top-down computations in word- and face-selective cortex. *eLife* 6:e22341
- Kim JS, Kanjlia S, Merabet LB, Bedny M. 2017. Development of the visual word form area requires visual experience: evidence from blind Braille readers. *J. Neurosci.* 37(47):11495–504
- Kliegl R, Grabner E, Rolfs M, Engbert R. 2004. Length, frequency, and predictability effects of words on eye movements in reading. *Eur. J. Cogn. Psychol.* 16(1–2):262–84
- Kolster H, Peeters R, Orban GA. 2010. The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. *J. Neurosci.* 30(29):9801–20
- Konkle T, Oliva A. 2012. A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74(6):1114–24
- Kubota EC, Joo SJ, Huber E, Yeatman JD. 2019. Word selectivity in high-level visual cortex and reading skill. *Dev. Cogn. Neurosci.* 36:100593
- Kwon M, Legge GE, Dubbels BR. 2007. Developmental changes in the visual span for reading. *Vis. Res.* 47(22):2889–900
- Kwon M, Liu R. 2019. Linkage between retinal ganglion cell density and the nonuniform spatial integration across the visual field. *PNAS* 116(9):3827–36
- Lamare M. 1892. Des mouvements des yeux dans la lecture. *Bull. Mém. Soc. Fr. Ophthalmol.* 10:354–64
- Langer N, Peysakhovich B, Zuk J, Drottar M, Sliva DD, et al. 2015. White matter alterations in infants at risk for developmental dyslexia. *Cereb. Cortex* 27(2):1027–36
- Latham K, Whitaker D. 1996. A comparison of word recognition and reading performance in foveal and peripheral vision. *Vis. Res.* 36(17):2665–74
- Le R, Witthoft N, Ben-Shachar M, Wandell B. 2017. The field of view available to the ventral occipitotemporal reading circuitry. *J. Vis.* 17(4):6
- Leehey SC, Cahn A. 1979. Lateral asymmetries in the recognition of words, familiar faces and unfamiliar faces. *Neuropsychologia* 17(6):619–28
- Legge GE, Cheung S-H, Yu D, Chung STL, Lee H-W, Owens DP. 2007. The case for the visual span as a sensory bottleneck in reading. *J. Vis.* 7(2):9
- Legge GE, Mansfield JS, Chung ST. 2001. Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vis. Res.* 41(6):725–43
- Lerma-Usabiaga G, Carreiras M, Paz-Alonso PM. 2018. Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *PNAS* 115(82):E9981–90
- Lescroart MD, Gallant JL. 2019. Human scene-selective areas represent 3D configurations of surfaces. *Neuron* 101(1):178–92.e7
- Levi DM. 2008. Crowding—an essential bottleneck for object recognition: a mini-review. *Vis. Res.* 48(5):635–54
- Levi DM, Klein SA, Aitsebaomo AP. 1985. Vernier acuity, crowding and cortical magnification. *Vis. Res.* 25(7):963–77
- Levy I, Hasson U, Avidan G, Hendler T, Malach R. 2001. Center-periphery organization of human object areas. *Nat. Neurosci.* 4(5):533–39
- Liu R, Patel BN, Kwon M. 2017. Age-related changes in crowding and reading speed. *Sci. Rep.* 7:8271
- Livingstone MS, Rosen GD, Drislane FW, Galaburda AM. 1991. Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *PNAS* 88(18):7943–47
- Livingstone MS, Vincent JL, Arcaro MJ, Srihasam K, Schade PF, Savage T. 2017. Development of the macaque face-patch system. *Nat. Commun.* 8:14897
- Lovegrove WJ, Bowling A, Badcock D, Blackwood M. 1980. Specific reading disability: differences in contrast sensitivity as a function of spatial frequency. *Science* 210(24):439–40
- Mackey WE, Winawer J, Curtis CE. 2017. Visual field map clusters in human frontoparietal cortex. *eLife* 6:e22974

- Mackworth NH. 1965. Visual noise causes tunnel vision. *Psychonom. Sci.* 3:67–68
- Malach R, Levy I, Hasson U. 2002. The topography of high-order human object areas. *Trends Cogn. Sci.* 6(4):176–84
- Manassi M, Whitney D. 2018. Multi-level crowding and the paradox of object recognition in clutter. *Curr. Biol.* 28(3):R127–33
- Marinkovic K, Dhond RP, Dale AM, Glessner M, Carr V, Halgren E. 2003. Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron* 38(3):487–97
- Martelli M, Di Filippo G, Spinelli D, Zoccolotti P. 2009. Crowding, reading, and developmental dyslexia. *J. Vis.* 9(4):14
- McCandliss BD, Cohen L, Dehaene S. 2003. The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* 7(7):293–99
- McClelland JL, Rumelhart DE. 1981. An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychol. Rev.* 88(5):375–407
- McConkie GW, Rayner K. 1975. The span of the effective stimulus during a fixation in reading. *Percept. Psychophys.* 17:578–86
- McKone E, Crookes K, Jeffery L, Dilks DD. 2012. A critical review of the development of face recognition: Experience is less important than previously believed. *Cogn. Neuropsychol.* 29(1–2):174–212
- Mirault J, Snell J, Grainger J. 2018. You that read wrong again! A transposed-word effect in grammaticality judgments. *Psychol. Sci.* 29(12):1922–29
- Mishkin M, Forgy DG. 1952. Word recognition as a function of retinal locus. *J. Exp. Psychol.* 43(1):43–48
- Mullin PA, Egeth HE. 1989. Capacity limitations in visual word processing. *J. Exp. Psychol. Hum. Percept. Perform.* 15(1):111–23
- Nandy AS, Sharpee TO, Reynolds JH, Mitchell JF. 2013. The fine structure of shape tuning in area V4. *Neuron* 78(6):1102–15
- Nasr S, Liu N, Devaney KJ, Yue X, Rajimehr R, et al. 2011. Scene-selective cortical regions in human and nonhuman primates. *J. Neurosci.* 31(39):13771–85
- Nazir TA, O'Regan JK, Jacobs AM. 1991. On words and their letters. *Bull. Psychon. Soc.* 29:171–74
- Nobre AC, Allison T, McCarthy G. 1994. Word recognition in the human inferior temporal lobe. *Nature* 372(6503):260–63
- Nordt M, Gomez J, Natu V, Rezai AA, Finzi D, et al. 2020. Cortical recycling in high-level visual cortex during childhood development. bioRxiv 209783. <https://doi.org/10.1101/2020.07.18.209783>
- O'Brien G, Yeatman J. 2021. Bridging sensory and language theories of dyslexia: toward a multifactorial model. *Dev. Sci.* 24(3):e13039
- Op de Beeck HP, Pillot I, Ritchie JB. 2019. Factors determining where category-selective areas emerge in visual cortex. *Trends Cogn. Sci.* 23(9):784–97
- Orbach J. 1952. Retinal locus as a factor in the recognition of visually perceived words. *Am. J. Psychol.* 65(4):555–62
- O'Regan JK, Jacobs AM. 1992. Optimal viewing position effect in word recognition: a challenge to current theory. *J. Exp. Psychol. Hum. Percept. Perform.* 18(1):185–97
- O'Regan JK, Lévy-Schoen A, Jacobs AM. 1983. The effect of visibility on eye-movement parameters in reading. *Percept. Psychophys.* 34(5):457–64
- Ossowski A, Behrmann M. 2015. Left hemisphere specialization for word reading potentially causes, rather than results from, a left lateralized bias for high spatial frequency visual information. *Cortex* 72:27–39
- Parvizi J, Jacques C, Foster BL, Witthoft N, Rangarajan V, et al. 2012. Electrical stimulation of human fusiform face-selective regions distorts face perception. *J. Neurosci.* 32(43):14915–20
- Pasupathy A, Popovkina DV, Kim T. 2020. Visual functions of primate area V4. *Annu. Rev. Vis. Sci.* 6:363–85
- Peelen MV, Downing PE. 2005. Selectivity for the human body in the fusiform gyrus. *J. Neurophysiol.* 93:603–8
- Pelli DG. 2008. Crowding: a cortical constraint on object recognition. *Curr. Opin. Neurobiol.* 18(4):445–51
- Pelli DG, Farell B, Moore DC. 2003. The remarkable inefficiency of word recognition. *Nature* 423(6941):752–56
- Pelli DG, Tillman KA. 2008. The uncrowded window of object recognition. *Nat. Neurosci.* 11(10):1129–35
- Pelli DG, Tillman KA, Freeman J, Su M, Berger TD, Majaj NJ. 2007. Crowding and eccentricity determine reading rate. *J. Vis.* 7(2):20

- Pennington BF. 2006. From single to multiple deficit models of developmental disorders. *Cognition* 101(2):385–413
- Pennington BF, Santerre-Lemmon L, Rosenberg J, MacDonald B, Boada R, et al. 2012. Individual prediction of dyslexia by single versus multiple deficit models. *J. Abnorm. Psychol.* 121(1):212–24
- Pestilli F, Carrasco M, Heeger DJ, Gardner JL. 2011. Attentional enhancement via selection and pooling of early sensory responses in human visual cortex. *Neuron* 72(5):832–46
- Pflugshaupt T, Gutbrod K, Wurtz P, von Wartburg R, Nyffeler T, et al. 2009. About the role of visual field defects in pure alexia. *Brain* 132(Pt. 7):1907–17
- Pollatsek A, Bolozky S, Well AD, Rayner K. 1981. Asymmetries in the perceptual span for Israeli readers. *Brain Lang.* 14(1):174–80
- Price CJ, Devlin JT. 2003. The myth of the visual word form area. *NeuroImage* 19(3):473–81
- Price CJ, Devlin JT. 2011. The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn. Sci.* 15(6):246–53
- Pugh KR, Landi N, Preston JL, Mencl WE, Austin AC, et al. 2013. The relationship between phonological and auditory processing and brain organization in beginning readers. *Brain Lang.* 125(2):173–83
- Radach R, Kennedy A. 2013. Eye movements in reading: some theoretical context. *Q. J. Exp. Psychol.* 66(3):429–52
- Rauschecker AM, Bowen RF, Parvizi J, Wandell BA. 2012. Position sensitivity in the visual word form area. *PNAS* 109(24):E1568–77
- Rayner K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124(3):372–422
- Rayner K, Schotter ER, Masson MEJ, Potter MC, Treiman R. 2016. So much to read, so little time: How do we read, and can speed reading help? *Psychol. Sci. Public Interest* 17(1):4–34
- Reich L, Szwed M, Cohen L, Amedi A. 2011. A ventral visual stream reading center independent of visual experience. *Curr. Biol.* 21(5):363–68
- Reicher GM. 1969. Perceptual recognition as a function of meaningfulness of stimulus material. *J. Exp. Psychol.* 81(2):275–80
- Reichle ED, Pollatsek A, Rayner K. 2006. E-Z Reader: a cognitive-control, serial-attention model of eye-movement behavior during reading. *Cogn. Syst. Res.* 7(1):4–22
- Reilly RG, Radach R. 2006. Some empirical tests of an interactive activation model of eye movement control in reading. *Cogn. Syst. Res.* 7(1):34–55
- Richlan F, Kronbichler M, Wimmer H. 2011. Meta-analyzing brain dysfunctions in dyslexic children and adults. *NeuroImage* 56(3):1735–42
- Riesenhuber M, Glezer LS. 2017. Evidence for rapid localist plasticity in the ventral visual stream: the example of words. *Lang. Cogn. Neurosci.* 32(3):286–94
- Riesenhuber M, Poggio T. 1999. Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2(11):1019–25
- Risse S. 2014. Effects of visual span on reading speed and parafoveal processing in eye movements during sentence reading. *J. Vis.* 14(8):11
- Rolfs M, Carrasco M. 2012. Rapid simultaneous enhancement of visual sensitivity and perceived contrast during saccade preparation. *J. Neurosci.* 32(40):13744–52a
- Rueckl JG, Paz-Alonso PM, Molfese PJ, Kuo W-J, Bick A, et al. 2015. Universal brain signature of proficient reading: evidence from four contrasting languages. *PNAS* 112(50):15510–15
- Salmelin R, Helenius P, Service E. 2000. Neurophysiology of fluent and impaired reading: a magnetoencephalographic approach. *J. Clin. Neurophysiol.* 17(2):163–74
- Saygin ZM, Osher DE, Norton ES, Youssofian DA, Beach SD, et al. 2016. Connectivity precedes function in the development of the visual word form area. *Nat. Neurosci.* 19(9):1250–55
- Scaltritti M, Balota DA. 2013. Are all letters really processed equally and in parallel? Further evidence of a robust first letter advantage. *Acta Psychol.* 144(2):397–410
- Schotter ER, Angele B, Rayner K. 2012. Parafoveal processing in reading. *Atten. Percept. Psychophys.* 74(1):5–35
- Seidenberg MS, McClelland JL. 1989. A distributed, developmental model of word recognition and naming. *Psychol. Rev.* 96(4):523–68

- Sereno SC, Rayner K. 2003. Measuring word recognition in reading: eye movements and event-related potentials. *Trends Cogn. Sci.* 7(11):489–93
- Sergent J. 1982. The cerebral balance of power: confrontation or cooperation? *J. Exp. Psychol. Hum. Percept. Perform.* 8(2):253–72
- Shepherdson P, Miller J. 2014. Redundancy gain in semantic categorisation. *Acta Psychol.* 148:96–106
- Siéoroff E, Haehnel-Benoliel N. 2015. Environmental script affects lateral asymmetry of word recognition: a study of French–Hebrew bilinguals tested in Israel and in France. *Laterality* 20(4):389–417
- Silson EH, Chan AW-Y, Reynolds RC, Kravitz DJ, Baker CI. 2015. A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *J. Neurosci.* 35(34):11921–35
- Silver MA, Kastner S. 2009. Topographic maps in human frontal and parietal cortex. *Trends Cogn. Sci.* 13(11):488–95
- Silver MA, Ress D, Heeger DJ. 2005. Topographic maps of visual spatial attention in human parietal cortex. *J. Neurophysiol.* 94(2):1358–71
- Snell J, Declerck M, Grainger J. 2018a. Parallel semantic processing in reading revisited: effects of translation equivalents in bilingual readers. *Lang. Cogn. Neurosci.* 33(5):563–74
- Snell J, Grainger J. 2017. The sentence superiority effect revisited. *Cognition* 168:217–21
- Snell J, Grainger J. 2019a. Consciousness is not key in the serial-versus-parallel debate. *Trends Cogn. Sci.* 23(10):814–15
- Snell J, Grainger J. 2019b. Readers are parallel processors. *Trends Cogn. Sci.* 23(7):537–46
- Snell J, Meeter M, Grainger J. 2017. Evidence for simultaneous syntactic processing of multiple words during reading. *PLOS ONE* 12(3):e0173720
- Snell J, van Leipsig S, Grainger J, Meeter M. 2018b. OBI-reader: a model of word recognition and eye movements in text reading. *Psychol. Rev.* 125(6):969–84
- Snowling M. 1998. Dyslexia as a phonological deficit: evidence and implications. *Child Psychol. Psychiatry Rev.* 3(1):4–11
- Song C, Schwarzkopf DS, Kanai R, Rees G. 2015. Neural population tuning links visual cortical anatomy to human visual perception. *Neuron* 85(3):641–56
- Squire LR, ed. 2009. *The History of Neuroscience in Autobiography*, Vol. 6. Oxford, UK: Oxford Univ. Press
- Srihasam K, Vincent JL, Livingstone MS. 2014. Novel domain formation reveals proto-architecture in inferotemporal cortex. *Nat. Neurosci.* 17(12):1776–83
- Stein J. 2001. The magnocellular theory of developmental dyslexia. *Dyslexia* 7(1):12–36
- Strasburger H. 2020. Seven myths on crowding and peripheral vision. *i-Perception* 11(3):2041669520913052
- Strother L, Coros AM, Vilis T. 2015. Visual cortical representation of whole words and hemifield-split word parts. *J. Cogn. Neurosci.* 28(3):194–98
- Strother L, Zhou Z, Coros AK, Vilis T. 2017. An fMRI study of visual hemifield integration and cerebral lateralization. *Neuropsychologia* 100:35–43
- Tadros K, Dupuis-Roy N, Fiset D, Arguin M, Gosselin F. 2013. Reading laterally: the cerebral hemispheric use of spatial frequencies in visual word recognition. *J. Vis.* 13(1):4
- Takemura H, Rokem A, Winawer J, Yeatman JD, Wandell BA, Pestilli F. 2016. A major human white matter pathway between dorsal and ventral visual cortex. *Cereb. Cortex* 26(5):2205–14
- Talcott JB, Witton C, Hebb GS, Stoodley CJ, Westwood EA, et al. 2002. On the relationship between dynamic visual and auditory processing and literacy skills; results from a large primary-school study. *Dyslexia* 8(4):204–25
- Tarr MJ, Cheng YD. 2003. Learning to see faces and objects. *Trends Cogn. Sci.* 7(1):23–30
- Tarr MJ, Gauthier I. 2000. FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nat. Neurosci.* 3(8):764–69
- Thesen T, McDonald CR, Carlson C, Doyle W, Cash S, et al. 2012. Sequential then interactive processing of letters and words in the left fusiform gyrus. *Nat. Commun.* 3:1284
- Toet A, Levi DM. 1992. The two-dimensional shape of spatial interaction zones in the parafovea. *Vis. Res.* 32(7):1349–57
- Tsao DY, Freiwald WA, Tootell RBH, Livingstone MS. 2006. A cortical region consisting entirely of face-selective cells. *Science* 311(5761):670–74

- Tydgat I, Grainger J. 2009. Serial position effects in the identification of letters, digits, and symbols. *J. Exp. Psychol. Hum. Percept. Perform.* 35(2):480–98
- Van der Haegen L, Cai Q, Brysbaert M. 2012. Colateralization of Broca's area and the visual word form area in left-handers: fMRI evidence. *Brain Lang.* 122(3):171–78
- Van der Haegen L, Cai Q, Seurinck R, Brysbaert M. 2011. Further fMRI validation of the visual half field technique as an indicator of language laterality: a large-group analysis. *Neuropsychologia* 49(10):2879–88
- Van der Haegen L, Cai Q, Stevens MA, Brysbaert M. 2013. Interhemispheric communication influences reading behavior. *J. Cogn. Neurosci.* 25(9):1442–52
- Vandermosten M, Vanderauwera J, Theys C, De Vos A, Vanvooren S, et al. 2015. A DTI tractography study in pre-readers at risk for dyslexia. *Dev. Cogn. Neurosci.* 14:8–15
- Vellutino FR, Fletcher JM, Snowling MJ, Scanlon DM. 2004. Specific reading disability (dyslexia): What have we learned in the past four decades? *J. Child Psychol. Psychiatry* 45(1):2–40
- Vidyasagar TR, Pammer K. 2010. Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends Cogn. Sci.* 14(2):57–63
- Vinckier F, Dehaene S, Jobert A, Dubus JP, Sigman M, Cohen L. 2007. Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron* 55(1):143–56
- Wade NJ, Tatler BW. 2008. Did Javal measure eye movements during reading? *J. Eye Mov. Res.* 2(5):5
- Wagner RK, Torgesen JK. 1987. The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol. Bull.* 101(2):192–212
- Wandell BA. 1995. *Foundations of Vision*. Sunderland, MA: Sinauer Assoc.
- Wandell BA, Dumoulin SO, Brewer AA. 2007. Visual field maps in human cortex. *Neuron* 56(2):366–83
- Wandell BA, Rauschecker AM, Yeatman JD. 2012. Learning to see words. *Annu. Rev. Psychol.* 63:31–53
- Wandell BA, Winawer J. 2010. Imaging retinotopic maps in the human brain. *Vis. Res.* 51(7):718–37
- Wandell BA, Winawer J. 2015. Computational neuroimaging and population receptive fields. *Trends Cogn. Sci.* 19(6):349–57
- Wandell BA, Yeatman JD. 2013. Biological development of reading circuits. *Curr. Opin. Neurobiol.* 23(2):261–68
- Weiner KS, Barnett MA, Lorenz S, Caspers J, Stigliani A, et al. 2017a. The cytoarchitecture of domain-specific regions in human high-level visual cortex. *Cereb. Cortex* 27(1):146–61
- Weiner KS, Barnett MA, Witthoft N, Golarai G, Stigliani A, et al. 2018. Defining the most probable location of the parahippocampal place area using cortex-based alignment and cross-validation. *NeuroImage* 170:373–84
- Weiner KS, Grill-Spector K. 2010. Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *NeuroImage* 52(4):1559–73
- Weiner KS, Grill-Spector K. 2011. Not one extrastriate body area: using anatomical landmarks, hMT+, and visual field maps to parcellate limb-selective activations in human lateral occipitotemporal cortex. *NeuroImage* 56(4):2183–99
- Weiner KS, Grill-Spector K. 2012. The improbable simplicity of the fusiform face area. *Trends Cogn. Sci.* 16(5):251–54
- Weiner KS, Yeatman JD, Wandell BA. 2017b. The posterior arcuate fasciculus and the vertical occipital fasciculus. *Cortex* 97:274–76
- Wen Y, Snell J, Grainger J. 2019. Parallel, cascaded, interactive processing of words during sentence reading. *Cognition* 189:221–26
- Wernicke C. 1881. *Lehrbuch der Gehirnkrankheiten für Aezerte und Studirende*. Kassel: Verl. Theodor Fischer
- Wernicke C. 1906. *Grundriss der Psychiatrie in klinischen Vorlesungen*. Leipzig: Thieme
- Wheat KL, Cornelissen PL, Sack AT, Schuhmann T, Goebel R, Blomert L. 2013. Charting the functional relevance of Broca's area for visual word recognition and picture naming in Dutch using fMRI-guided TMS. *Brain Lang.* 125(2):223–30
- Wheeler DD. 1970. Processes in word recognition. *Cogn. Psychol.* 1(1):59–85
- White AL, Boynton GM, Yeatman JD. 2019a. The link between reading ability and visual spatial attention across development. *Cortex* 121:44–59
- White AL, Boynton GM, Yeatman JD. 2019b. You can't recognize two words simultaneously. *Trends Cogn. Sci.* 23(10):812–14

- White AL, Palmer J, Boynton GM. 2018. Evidence of serial processing in visual word recognition. *Psychol. Sci.* 29(7):1062–71
- White AL, Palmer J, Boynton GM. 2020. Visual word recognition: evidence for a serial bottleneck in lexical access. *Atten. Percept. Psychophys.* 82(4):2000–17
- White AL, Palmer J, Boynton GM, Yeatman JD. 2019c. Parallel spatial channels converge at a bottleneck in anterior word-selective cortex. *PNAS* 116(20):10087–96
- White AL, Runeson E, Palmer J, Ernst ZR, Boynton GM. 2017. Evidence for unlimited capacity processing of simple features in visual cortex. *J. Vis.* 17(6):19
- Whitney C. 2001. How the brain encodes the order of letters in a printed word: the SERIOL model and selective literature review. *Psychonom. Bull. Rev.* 8:221–43
- Whitney C. 2008. Supporting the serial in the SERIOL model. *Lang. Cogn. Process.* 23(6):824–65
- Whitney C, Lavidor M. 2004. Why word length only matters in the left visual field. *Neuropsychologia* 42(12):1680–88
- Winawer J, Horiguchi H, Sayres RA, Amano K, Wandell BA. 2010. Mapping hV4 and ventral occipital cortex: the venous eclipse. *J. Vis.* 10(5):1
- Wolf M, Bowers PG. 1999. The double-deficit hypothesis for the developmental dyslexias. *J. Educ. Psychol.* 91(3):415–38
- Wong AC-N, Jobard G, James KH, James TW, Gauthier I. 2009. Expertise with characters in alphabetic and nonalphabetic writing systems engage overlapping occipito-temporal areas. *Cogn. Neuropsychol.* 26(1):111–27
- Woodhead ZVJ, Barnes GR, Penny W, Moran R, Teki S, et al. 2014. Reading front to back: MEG evidence for early feedback effects during word recognition. *Cereb. Cortex* 24(3):817–25
- Woolnough O, Donos C, Rollo PS, Forseth KJ, Lakretz Y, et al. 2021. Spatiotemporal dynamics of orthographic and lexical processing in the ventral visual pathway. *Nat. Hum. Behav.* 5:389–98
- Yeatman JD, Dougherty RF, Ben-Shachar M, Wandell BA. 2012. Development of white matter and reading skills. *PNAS* 109(44):E3045–53
- Yeatman JD, Rauschecker AM, Wandell BA. 2013. Anatomy of the visual word form area: adjacent cortical circuits and long-range white matter connections. *Brain Lang.* 125(2):146–55
- Yeatman JD, Weiner KS, Pestilli F, Rokem A, Mezer A, Wandell BA. 2014. The vertical occipital fasciculus: a century of controversy resolved by in vivo measurements. *PNAS* 111(48):E5214–23
- Yeh F-C, Panesar S, Fernandes D, Meola A, Yoshino M, et al. 2018. Population-averaged atlas of the macroscale human structural connectome and its network topology. *NeuroImage* 178:57–68
- Yu D, Cheung S-H, Legge GE, Chung STL. 2007. Effect of letter spacing on visual span and reading speed. *J. Vis.* 7(2):2
- Yu D, Legge GE, Wagoner G, Chung STL. 2018. Training peripheral vision to read: boosting the speed of letter processing. *Vis. Res.* 152:51–60
- Zeki S. 2016. Multiple asynchronous stimulus- and task-dependent hierarchies (STDH) within the visual brain's parallel processing systems. *Eur. J. Neurosci.* 44(8):2515–27



Contents

Coming of Age in Science: Just Look? <i>Ken Nakayama</i>	1
The Diversity of Eyes and Vision <i>Dan-E. Nilsson</i>	19
Generating and Using Transcriptomically Based Retinal Cell Atlases <i>Karthik Shekhar and Joshua R. Sanes</i>	43
Morphology, Molecular Characterization, and Connections of Ganglion Cells in Primate Retina <i>Ulrike Grünert and Paul R. Martin</i>	73
Impact of Photoreceptor Loss on Retinal Circuitry <i>Joo Yeun Lee, Rachel A. Care, Luca Della Santina, and Felice A. Dunn</i>	105
Imaging the Retinal Vasculature <i>Stephen A. Burns, Ann E. Elsner, and Thomas J. Gast</i>	129
The Mechanism of Macular Sparing <i>Jonathan C. Horton, John R. Economides, and Daniel L. Adams</i>	155
Neuromodulatory Control of Early Visual Processing in Macaque <i>Anita A. Disney</i>	181
Visual Signals in the Mammalian Auditory System <i>Meredith N. Schmebl and Jennifer M. Grob</i>	201
Population Models, Not Analyses, of Human Neuroscience Measurements <i>Justin L. Gardner and Elisha P. Merriam</i>	225
Visual Remapping <i>Julie D. Golomb and James A. Mazer</i>	257
Role of the Superior Colliculus in Guiding Movements Not Made by the Eyes <i>Bonnie Cooper and Robert M. McPeck</i>	279

Spatial Integration in Normal Face Processing and Its Breakdown in Congenital Prosopagnosia <i>Galia Avidan and Marlene Bebrmann</i>	301
Urgent Decision Making: Resolving Visuomotor Interactions at High Temporal Resolution <i>Terrence R. Stanford and Emilio Salinas</i>	323
Remembering the Past to See the Future <i>Nicole C. Rust and Stephanie E. Palmer</i>	349
Balancing Flexibility and Interference in Working Memory <i>Timothy J. Buschman</i>	367
Binocular Vision and Stereopsis Across the Animal Kingdom <i>Jenny C.A. Read</i>	389
Lightness Perception in Complex Scenes <i>Richard F. Murray</i>	417
Quantifying Visual Image Quality: A Bayesian View <i>Zhengfang Duanmu, Wentao Liu, Zhongling Wang, and Zhou Wang</i>	437
The Certainty of Ambiguity in Visual Neural Representations <i>Jan W. Brascamp and Steven K. Shevell</i>	465
Reading: The Confluence of Vision and Language <i>Jason D. Yeatman and Alex L. White</i>	487
Eyewitness Identification Is a Visual Search Task <i>John T. Wixted, Edward Vul, Laura Mickes, and Brent M. Wilson</i>	519
Face Recognition by Humans and Machines: Three Fundamental Advances from Deep Learning <i>Alice J. O'Toole and Carlos D. Castillo</i>	543
Mobile Computational Photography: A Tour <i>Mauricio Delbracio, Damien Kelly, Michael S. Brown, and Peyman Milanfar</i>	571
Lexical Color Categories <i>Dehwin T. Lindsey and Angela M. Brown</i>	605
Aging of the Retina: Molecular and Metabolic Turbulences and Potential Interventions <i>Laura Campello, Nivedita Singh, Jaysbree Advani, Anupam K. Mondal, Ximena Corso-Díaz, and Anand Swaroop</i>	633
Retina Metabolism and Metabolism in the Pigmented Epithelium: A Busy Intersection <i>James B. Hurley</i>	665

Optical Coherence Tomography and Glaucoma <i>Alexi Geevarghese, Gadi Wollstein, Hiroshi Ishikawa, and Joel S. Schuman</i>	693
Genetic Determinants of Intraocular Pressure <i>Zihe Xu, Pirro Hysi, and Anthony P. Khawaja</i>	727
Measures of Function and Structure to Determine Phenotypic Features, Natural History, and Treatment Outcomes in Inherited Retinal Diseases <i>Artur V. Cideciyan, Arun K. Krishnan, Alejandro J. Roman, Alexander Sumaroka, Malgorzata Swider, and Samuel G. Jacobson</i>	747
Eye Movements in Macular Degeneration <i>Preeti Verghese, Cécile Vullings, and Natela Shamidze</i>	773
Functional Organization of Extraocular Motoneurons and Eye Muscles <i>Anja K.E. Horn and Hans Straka</i>	793
Axonal Growth Abnormalities Underlying Ocular Cranial Nerve Disorders <i>Mary C. Whitman</i>	827
Precision Medicine Trials in Retinal Degenerations <i>Sarah R. Levi, Joseph Ryu, Pei-Kang Liu, and Stephen H. Tsang</i>	851
Best Practices for the Design of Clinical Trials Related to the Visual System <i>Maureen G. Maguire</i>	867

Errata

An online log of corrections to *Annual Review of Vision Science* articles may be found at <http://www.annualreviews.org/errata/vision>