Neuropsychologia 95 (2017) 136–155

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Review article

Why are digits easier to identify than letters?

Teresa M. Schubert

ARC Centre of Excellence in Cognition and its Disorders (CCD), Department of Cognitive Science, Macquarie University, 16 University Ave., Sydney, NSW 2109, Australia

A R T I C L E I N F O

Keywords:
Letter identification
Digit identification
Acquired dyslexia
Reading
Numerical processing

A B S T R A C T

Beginning with Dejerine's report of pure alexia in 1892, numerous researchers have noted that individuals with acquired impairments of reading may show spared digit identification performance. This digit advantage has also been found in unimpaired adult readers across a number of tasks, and five main hypotheses have been proposed to explain how it arises. In this paper I consider these hypotheses in the context of recent theories of a unified alphanumeric character identification system, and evaluate them according to relevant empirical evidence. Despite some promising findings, none of the hypotheses currently provide a sufficient explanation of the digit advantage. Rather than developing new hypotheses to explain a categorical difference between digit and letter performance, I argue that future work should consider factors that affect identification performance specific to individual characters.

1. Introduction

Much recent research has focused on the processing of letters and words, which form a major component of our daily lives. Less work has considered Arabic digits (0–9), which are also prevalent visual stimuli that we process with ease. As visual stimuli, digits and letters are fairly similar in form and consensus is emerging that identification processes are shared between the two character types (Grainger and Hannagan, 2014; Kinoshita and Lagoutaris, 2010; McCloskey and Schubert, 2014). These theories are backed by a growing body of evidence from normal and impaired readers supporting the similarity of digit and letter identification. In the face of these data and corresponding theories, it has also been noted that digit identification is often more accurate than letter identification, and many authors have proposed possible explanations for this phenomenon (Cohen and Dehaene, 1995; Holender and Peereman, 1987; Ingles and Eskes, 2008; Polk et al., 2002; Polk and Farah, 1998; Rath et al., 2015; Starrfelt and Behrmann, 2011). However, few of these explanations have been explored or tested. In this paper I consider hypotheses for the digit identification advantage, evaluating them with respect to relevant properties of the characters and within an alphanumeric character identification system.

1.1. Evidence for shared letter/digit processing

Evidence has accumulated over the past few decades for similarities of performance in letter and digit identification tasks. This evidence has come not just from unimpaired adult readers but also acquired and developmental dyslexia. From adult readers, the evidence is largely from partial report/Reicher-Wheeler tasks in which strings of random letters, digits, or non-alphanumeric symbols are presented (e.g., Collis et al., 2013; Hammond and Green, 1982; Tydgat and Grainger, 2009). After stimulus offset, participants report whether a probe letter was present, or report the letter in a post-cued position; either version of the task requires identity and/or position processing for the characters of the string. A large number of studies have found that performance (indexed by the shape of the serial position function) is similar for letter and digit strings (Chanceaux and Grainger, 2012; Collis et al., 2013; Duñabeitia et al., 2012; García-Orza et al., 2010; Hammond and Green, 1982; Tydgat and Grainger, 2009), and only one early study reported a difference between the two character types (Mason, 1982).

Single letter and digit identification in unimpaired adults was studied by Starrfelt et al. (2010). Adult participants were asked to name a single briefly presented and masked character (from the digits 0–9 and uppercase letters A-J). The characters were blocked by type and presented in random order. Performance (as visual processing speed) was found to be approximately equivalent for digits and letters (Starrfelt et al., 2010). Other researchers have used priming in the same/different match task and lexical decision task to demonstrate that letters and digits activate the same sets of visual features and show effects of visual similarity across character types, providing further evidence for shared identification processing (Kinoshita and Lagoutaris, 2010; Kinoshita et al., 2015; Kinoshita et al., 2013; Perea et al., 2008).

Similarities between letter and digit identification processing are

E-mail address: teresa.schubert@mq.edu.au.

http://dx.doi.org/10.1016/j.neuropsychologia.2016.12.016
Received 23 September 2016; Received in revised form 25 November 2016; Accepted 11 December 2016
Available online 13 December 2016
0028-3932/

also seen in individuals with dyslexia. There have been three reports of individuals with acquired dyslexia who have highly similar deficits in identifying both character types (Katz and Sevush, 1989; McCloskey and Schubert, 2014; Patterson and Wilson, 1990). The similarities concern identification error types, serial position functions, as well as cross-category substitution errors. Furthermore, in a sample of five individuals with pure alexia studied by Starrfelt and colleagues (Starrfelt and Behrmann, 2011; Starrfelt et al., 2010, 2009), all showed impaired performance relative to controls on letter and digit identification tasks (four also showed numerically lower performance with letters than digits). In developmental dyslexia, there has been one report of a severe letter-identification deficit which also affected digits (Brunsdon et al., 2006).[1] Furthermore, a large group study of children with developmental dyslexia performing a partial report task also reported comparable performance for letter and digit stimuli (Ziegler et al., 2010), as did a study of adults with developmental dyslexia (Collis et al., 2013).

1.1.1. Shared letter and digit identification theories

The convergence of data from various lines of research onto the conclusion that letters and digits share an identification system is highly persuasive, and three research groups have posited theories with this property. The earliest theory is by Norris, Kinoshita, and colleagues (Norris et al., 2010; Norris and Kinoshita, 2012), who describe a model of letter identification as an instance of general object recognition. Kinoshita and Lagoutaris (2010) made a more specific claim: Letter and digit stimuli directly compete, activating both letter and digit identities due to their shared visual features. No distinction is drawn between letter and digit identification in this system. The second theory of a shared alphanumeric identification system, posited by McCloskey and Schubert (2014), includes a level at which visual features are represented, a level at which stored visual forms (allographs) are represented, and finally a level of character identities, with concurrent access to category information. Though category information is available within the system, the authors posit that it does not affect identity processing. Finally, Grainger and colleagues (Grainger et al., 2016b; Grainger and Hannagan, 2014) also suggest that letters and digits are recognized by the same process, contacting position-dependent character identities ('alphanumeric detectors'). By contrast to the other models, Grainger et al. (2016a) posit that letter and digit processing diverges prior to any position-invariant representations of character identity. However, the earliest stage of the model makes no distinction between letter and digit stimuli.

Though the details of these three theories differ, the basic assumption of a shared identification system is present in all of them: Digits and letters are identified in the same system without distinction based on the category of the stimulus. One major finding that seems to challenge shared alphanumeric identification is the alphanumeric category effect in visual search. This effect refers to searches for a different-category target (e.g., digit among letters) being more efficient than searches for a same-category target (e.g., digit among digits) (e.g., Egeth et al., 1972; Jonides and Gleitman, 1972; Polk and Farah, 1998; Taylor, 1978). The alphanumeric category effect has been taken as evidence for an at least partially segregated character identification system (Hamilton et al., 2006; Polk and Farah, 1998). However, the alphanumeric category effect can be accounted for without separate letter and digit identification by positing that category information is extracted in parallel with identity information (McCloskey and Schubert, 2014; Taylor, 1978).

Given the premise of shared identification processing without regard to category, it would be simplest to assume that digit and letter identification would be performed with equal accuracy and speed. However, this is not necessarily the case because identification may depend on characteristics such as frequency of occurrence, visual similarity of the stimulus to other characters, and the influences of downstream processing via feedback. The shared alphanumeric identification theories have not been implemented to the level of comparing letter and digit identification accuracy. This is in part due to a lack of knowledge about the effects of these variables, but also reflects the difficulty of modelling the full range of human identification performance for letters and digits across size, case, font, handwriting style, and other sources of variability in the input (e.g., Chang et al., 2012; Finkbeiner and Coltheart, 2009). In fact, empirical evidence suggests that digits often enjoy a speed or accuracy advantage in identification.

1.2. The digit identification advantage

Though a large body of work, reviewed above, suggests that letter and digit identification are similar, other findings suggest that digits may be easier to identify in some contexts. Individuals with dyslexia as well as unimpaired readers have been found to identify single digits faster and/or more accurately than single letters in a variety of tasks.

The main body of evidence for the digit identification advantage is from cases of acquired reading impairment. Possibly the first evidence was reported in 1892 by Dejerine (as translated and discussed by Bub et al. (1993)); the individual he described was poor at recognizing single letters but succeeded at recognizing single digits. A similar observation is commonly reported in studies of acquired dyslexia, where letters are often affected more severely than digits (e.g., Cohen and Dehaene, 1995; Deloche and Seron, 1987; Greenblatt, 1973; Grossi et al., 1984; Ingles and Eskes, 2008; Larsen et al., 2004; McCloskey and Schubert, 2014; Perri et al., 1996). In a review of 90 cases of pure alexia, Starrfelt and Behrmann (2011) reported that these individuals generally have an impairment in both single letter and single digit processing, though letters tend to be more impaired. They found no dissociations in which a clear digit identification impairment was found in the face of intact letter identification. And finally, a recent paper by Rath and colleagues (Rath et al., 2015) confirms that digit naming impairment with intact letter naming has not been reported in the aphasia literature. They also present new evidence that an advantage for digit processing over letter/word reading was present in a large unselected sample of individuals with aphasia (Rath et al., 2015).

According to these studies, and to my knowledge, there have been no reported individuals with impaired digit identification in the face of intact letter identification. However, there are a few cases of digit naming impairment which may be instructive. For example, Cipolotti (1995) report the case of SF, who was severely impaired in naming multi-digit numbers, but showed normal reading performance. It is interesting to note that only a small proportion (14%) of SF's errors in number naming were classified as lexical errors (e.g., 54 as 'thirty-four'); the majority were syntactic errors (e.g., 54 as 'forty-five') or combination errors (e.g., 54 as 'forty-three'). Lexical errors could arise due to misidentification of the digits in the stimulus, while syntactic errors reflect correct identification of the digits but a failure in constructing the appropriate syntactic frame for the verbal number response (Dehaene, 1992; McCloskey, 1992). The combination of intact letter identification (as reading) and a majority of digit errors which are not based in identification suggests that selective deficits to digit processing arise after identification, and hence after letter and digit processing have diverged.

The digit identification advantage has also been found in adults without reading impairment, typically as a speed advantage. Ingles and Eskes (2008) compared letter and digit identification performance of one individual with acquired dyslexia to five control participants with brain damage but unimpaired reading. All of these participants completed an attentional blink task requiring identification of two

[1] Though developmental dyslexia and developmental dyscalculia have been found to dissociate (Butterworth, 2005; Landerl et al., 2009), neither generally entails a deficit in letter or digit identification, except in particular cases discussed here (e.g., Brunsdon et al. (2006) and possibly: Shalev and Gross-Tsur (1993)).

target digits or upper case letters (trials blocked by target type) at varying stimulus onset asynchronies (SOAs). At the only SOA at which control subjects were below ceiling for identification of the two targets, they showed an advantage in identifying digits over letters. The acquired dyslexic individual was at ceiling for identifying the first target when it was a digit, but had lower accuracy than controls when the first target was a digit. He showed lower accuracy than controls for both letters and digits as the second target (Ingles and Eskes, 2008). The commonality across these results is better performance with digits than letters, in both speed (SOA length) and accuracy. Starrfelt and Behrmann (2011) used the same single character naming task as Starrfelt et al. (2010) with a sample of young adults. At all durations that did not result in ceiling performance, average digit accuracy was superior to average letter accuracy. The authors suggested that the digit accuracy advantage often reported with impaired participants might be an amplification of a subtle digit speed advantage present in unimpaired participants (Starrfelt and Behrmann, 2011).

Digit and letter identification is largely trivial for unimpaired adult participants, which suggests that when a digit advantage has not been found it could be masked by ceiling-level identification performance. For example, the Ingles and Eskes (2008), and Starrfelt and Behrmann (2011) studies found no difference between letter and digit performance in the easier task conditions (longer SOA, longer exposure duration). As far as I know, no one has directly compared performance on strings of letters and digits in unimpaired adults. Thus it is unclear whether a digit advantage in unimpaired participants is universal (given sufficient task difficulty), or whether it arises only in single character identification tasks, like those reviewed in this section.

The digit advantage has been reported across populations (impaired and unimpaired readers) and relative to both uppercase and lowercase letters. Most studies have found it in tasks with blocked presentation of the letters and digits (e.g., Grossi et al., 1984; Larsen et al., 2004; Perri et al., 1996; Starrfelt et al., 2010, 2009), but one study also found it in mixed presentation in a case of acquired dyslexia (McCloskey and Schubert, 2014). And finally, it seems to arise when the entire alphabet is presented (e.g., Cohen and Dehaene, 1995; McCloskey and Schubert, 2014) and also when just a subset is used (e.g., Ingles and Eskes, 2008; Starrfelt and Behrmann, 2011; Starrfelt et al., 2010, 2009). Unfortunately, the vast majority of studies do not report sufficient detail to determine all of these properties. Despite the unknowns about many studies, the details of some studies help constrain the source of the advantage. As will be discussed in more detail for the relevant hypotheses, some hypotheses may not be compatible with all of the digit advantage data based on these properties.

The lack of a double dissociation between letter and digit identification suggests that letter and digit processing may not be fully independent, and the consistent direction of the single dissociation (letters more impaired than digits) is intriguing. The accuracy advantage for digit over letter identification might be considered evidence for separate identification systems, with the digit system operating more quickly or efficiently in some manner. However, even within the set of letters, some letters are identified more quickly than others, including by unimpaired readers (e.g., Fiset et al., 2008; Jones and Mewhort, 2004; Mueller and Weidemann, 2012; Pitchford et al., 2008), but this is not considered evidence for separate identification systems for different letters. Given the converging evidence across impaired and unimpaired populations for a shared system, it would be preferable to explain the digit advantage within a single identification system, rather than return to separate systems. Across previous studies reporting the digit advantage, authors have suggested multiple hypotheses to explain the digit advantage (Cohen and Dehaene, 1995; Holender and Peereman, 1987; Ingles and Eskes, 2008; Polk et al., 2002; Polk and Farah, 1998; Rath et al., 2015; Starrfelt and Behrmann, 2011); the remainder of this paper discusses these hypotheses in the context of shared alphanumeric identification.

2. What produces the digit advantage?

Five possible sources of the digit advantage have been discussed in the literature, generally in the context of case reports of acquired identification impairments. These hypotheses are: (1) differences in visual properties (Cohen and Dehaene, 1995; Polk et al., 2002; Starrfelt and Behrmann, 2011), (2) differences in character frequency (Ingles and Eskes, 2008; Rath et al., 2015), (3) the smaller set of digit identities (Cohen and Dehaene, 1995; Ingles and Eskes, 2008; Polk et al., 2002; Polk and Farah, 1998), (4) a boost from number semantics (Cohen and Dehaene, 1995; Ingles and Eskes, 2008; Rath et al., 2015; Starrfelt and Behrmann, 2011), and (5) a supporting role for the right hemisphere (Cohen and Dehaene, 1995; Holender and Peereman, 1987). Confirmation of one or more of these hypotheses would reduce any perceived tension between the reported digit advantage and an underlying shared identification system. Furthermore, it would enrich our understanding of how the identification system functions by delineating properties of letters and digits relevant to the identification process.

The viability of the first two hypotheses (differences in visual properties or frequency) depends on whether their preconditions are true. For example, a difference between visual similarity of digits and letters can only explain the digit advantage if such a difference exists, and further if it is in the direction favouring digit identification. In the next few sections I evaluate the preconditions of these hypotheses by comparing digits and letters on a number of metrics. The final three hypotheses (number of identities, semantics, right hemisphere) are currently harder to evaluate, but I discuss their viability in a shared identification system (McCloskey and Schubert, 2014), and sources of evidence that might be relevant. As the final two hypotheses are not always distinct, and are often discussed in relation to the neural bases of number processing, I discuss them in conjunction.

2.1. Hypothesis 1: digits are simpler or more distinctive visual forms

In this section I consider whether there are systematic differences between letters and digits in visual properties. Full support for the hypothesis that a particular property (e.g., pixel overlap) is responsible for the digit advantage requires not only the precondition to be true (e.g., digits have a smaller degree of pixel overlap than letters) but also evidence that this difference affects identification performance. Instances in which a difference arises would invite empirical studies to determine whether the observed difference impacts character identification.

When first suggesting that visual properties might underlie the digit advantage, Cohen and Dehaene (1995) suggested that if letters were more visually similar to other letters than digits to other digits, it could be easier to identify a digit from among its competitors than a letter from its competitors (see also Polk et al., 2002). Their suggestion is agnostic about the particular visual property that might be relevant, but visual properties investigated for their role in letter identification include: degree of visual overlap among letters, visual discriminability of a given letter from other letters, and visual complexity (e.g, Arguin et al., 2002; Mueller and Weidemann, 2012; Pelli et al., 2006; Simpson et al., 2012; Starrfelt and Behrmann, 2011; Starrfelt et al., 2015).

Starrfelt and Behrmann (2011) were the first to explicitly test for the presence of a visual discriminability difference between letters and digits. Discriminability refers to how easy it is to visually discriminate a character from others, i.e., how dissimilar it is to other letters. For (dis) similarity they computed the mean pairwise pixel overlap for lowercase letters and digits in two common fonts (Times and Arial). Using this measure, they computed the number of competitors for letter and digit stimuli: the number of same-category stimuli with an overlap value higher than a set threshold. At all thresholds they considered, digits had the same number or fewer competitors than letters, and therefore they concluded that digits have higher discriminability than letters (Starrfelt and Behrmann, 2011). (At the lowest threshold this is necessarily true because there are fewer digits than letters).

This result suggests that digits are easier to pick out from the set of digits than letters from the set of letters. However, this explanation assumes that digit and letter identities are being selected from the set of same-category identities: digits among digits and letters among letters. Within the proposed shared identification systems described above, this assumption is not straightforward. In some situations high-level knowledge elsewhere in the reader's cognitive system, such as a checking or monitoring process (conscious or unconscious) might determine whether the selected letter identity matches the expected category; for example, when reading a telephone number one expects only digits. This covers pure-category contexts, when only digits or letters are expected due to external knowledge, or only digits or letters are boosted by their recent activation. However, pure-category contexts are not the only ones in which the accuracy discrepancy between letters and digits has been found. For this reason – to produce a more robust test for a discriminability effect that would generalise across testing contexts – I compute a measure that takes into account competitors from both categories.

I compute and compare letter and digit visual overlap, visual discriminability, and visual complexity separately, testing multiple measures to achieve a broad picture of the visual properties of letter and digit forms. For visual discriminability, it is first necessary to define the visual forms of the characters, which was done using both pixel overlap and feature overlap. While pixel overlap has been used to evaluate visual similarity and discriminability (Marinus et al., 2016; Starrfelt and Behrmann, 2011; Wong and Szücs, 2013), it is not a measure of abstract visual similarity because it is dependent on the font (typeface) used for the stimuli. This sensitivity to low-level visual attributes (e.g., serifs) may not accurately reflect the level(s) of representation at which letters and digits are perceptually similar (see related arguments in Starrfelt et al., 2015). Furthermore, pixel overlap measures fail to acknowledge that our experience with letters and digits spans more than a single font. The digit advantage is unlikely to be font specific, since it has been reported across a range of studies using an unknown variety of fonts for letter and digit stimuli. Therefore, visual similarity measures that index somewhat abstract letter features (e.g., horizontal bar, curve opened left), proposed to represent visual information at a higher perceptual level, may provide a more valid visual similarity metric. To this end, I constructed a feature set to describe basic letter and digit shapes. Throughout this section I consider uppercase letters, which provides a more conservative test for visual differences [2] and simplifies the feature set required to describe the letter forms. Furthermore, all of the studies that report the digit advantage and indicate the case of their letter stimuli used uppercase letters.

2.1.1. The fonts

Four common fonts were used for the pixel overlap measures: Arial, Consolas, Courier New, and Times New Roman (Fig. 1). These fonts were chosen to provide coverage of two dimensions of typography: serif presence/absence and fixed/proportional width. Consolas and Arial are sans serif fonts, while Courier New and Times New Roman are serif fonts. Consolas and Courier New are fixed width fonts, while Times New Roman and Arial are proportional fonts. All of these fonts are reasonably prevalent in printed material; Arial is particularly common for text displayed on a computer screen (Bernard et al., 2003; Moret-Tatay and Perea, 2011).

ABCDEFGHIJKLMNOPQRSTUVWXYZ
0123456789

ABCDEFGHIJKLMNOPQRSTUVWXYZ
0123456789

ABCDEFGHIJKLMNOPQRSTUVWXYZ
0123456789

ABCDEFGHIJKLMNOPQRSTUVWXYZ
0123456789

Fig. 1. The four fonts used in this study: Consolas (sans serif, fixed-width), Arial (sans serif, proportional), Courier New (serif, fixed-width), Times New Roman (serif, proportional), respectively. Though the fonts differ in physical size when reproduced at the same font size, this difference is irrelevant because comparisons were always conducted within font.

2.1.2. The feature set

The feature set developed here was meant to approximate the type of abstract feature representations hypothesized at the character shape level (Caramazza and Hillis, 1990; McCloskey and Schubert, 2014; Schubert and McCloskey, 2013). Much prior work has aimed to determine the features or sub-feature elements (e.g., line terminations, intersections/vertices) that are used in identification. Despite a large number of previous proposals for features that are relevant in letter identification, few proposed feature sets are specified to a level that allows computation of feature overlap; for example, most do not consider position or relative size. None of the existing feature sets describe digits. Guidance for the feature set described here came from these previous proposals (e.g., Boles and Clifford, 1989; Briggs and Hocevar, 1975; Chang et al., 2012; Changizi and Shimojo, 2005; Fiset et al., 2008; Lanthier et al., 2009; Pelli et al., 2006; Petit and Grainger, 2002; Rosa et al., 2016).

Some attempts to determine the relevant features in letter and word identification have used particular font forms (e.g., serif letters in Rosa et al. (2016)) but descriptions of a particular font may not reflect abstract character features. Accordingly, I intentionally do not account for the full range of font variability, which tends to be systematic alterations of the basic letter shapes (through changes of e.g., serifs, x-height, letter slant Sanocki and Dyson, 2012). In the terms of Hofstadter and McGraw (1995), the feature set describes the character "conceptualization" rather than the "letterform" of a specific font. In cases of possible ambiguity about the standard form, a sans-serif and fixed-width font (Consolas) served as a basis for the letter and digit feature descriptions.[3]

The features include lines and curves, which are assumed to be invariant to character size and position in the visual field. However, they can differ in relative size (i.e., one feature can be larger than another) and position within a character (e.g., feature 'horizontal line of length 1 at top of character' is distinguished from 'horizontal line of length 1 at bottom of character'). In feature sets proposed previously, letters have been split into features at points of discontinuity in the

[2] Upper case letters occupy the same region of space on a line as digits, unlike lowercase letters that often occupy a smaller region [e.g., x], or extend below the baseline [e.g., j]. The digit shapes are quite similar in that they all occupy the same region of space; comparing this degree of similarity to uppercase letters, which have the same property, increases the measured overlap between letters and digits and reduces the likelihood of finding a difference between the two types.

[3] The only exception is the zero character, which in Consolas is a crossed-zero (0). This is a particularly unusual form of this character and therefore the standard form (0) was described in the feature set.

visual stimulus (Changizi and Shimojo, 2005). In the interest of using a small set of atomic feature units to describe many letters, some letters were split into smaller segments in the absence of a discontinuity in the line. For example, U was split into two straight segments and one curve, whereas Changizi and Shimojo treat it as single stroke. This division of the character captures that U contains vertical segments which are quite similar to those in H, D, etc. It also avoids the use of whole-character features, which would violate findings that whole-letter templates are not used in identification (Pelli et al., 2006). A full listing of the features is in Appendix A.

In computing overlap between two characters, highest overlap was given for identical features at the same position (i.e., matching position, size, feature type, and feature class), and reduced overlap was given when some but not all of these aspects are shared (e.g., 1 of 3 possible points of overlap are given when feature types match despite differences in size and position). This captures the assumption that there is similarity in the representation of features that differ only in these parameters, and also provides a less sparse overlap matrix than would be obtained if only perfect matches were considered overlapping. See Table A2 for full details of the procedure for computing feature overlap.

The feature set combined with the overlap metric captures the intuitive similarity among characters, for example, giving high overlap to 1 and T, and 6 and 9. However, it is untested with regards the actual features used by the visual system. To the extent that the feature set is based on prior proposals for letter feature sets (themselves based on different types of empirical evidence) it should approximate features used in identification. A good test of the feature overlap metric would be relating it to identification performance, however there are no studies providing the relevant performance: by-character accuracy (or RT) for all uppercase letters and digits. As will be discussed in Section 3.2., this type of data is key to further our understanding of the character identification system.

2.1.3. Visual overlap

The average visual overlap among the set of digits was compared to the average for the set of letters, for both pixel and feature measures. Overlap was calculated for every pairwise combination of characters (1296 combinations), and separately for each font for pixel overlap. Subsequently, for each character, overlap values for the comparison between that character and each other character were averaged. Finally, the average overlap values for the 10 digits were compared to the average overlap values for the 26 letters. Throughout this paper, comparisons were conducted using nonparametric statistics (Spearman's rank-order correlation and Mann-Whitney U-test) and two-tailed significance testing. These techniques are well suited to the unequal variances and small and unequal sample sizes (i.e., N_1=10 digits, N_2=26 letters). The overlap values for each character (as well as all other visual characteristics compared in subsequent sections) are given in Appendix B.

2.1.3.1. Overlap: pixel. The pixel overlap metric was computed in MATLAB2013a (Mathworks) by computing the overlap of black pixels in images of single letter and digit stimuli. Characters were in 36 pt font and centred (within 300×300 pixels of white background) for maximal overlap (Marinus et al., 2016).

Comparisons were made within each font. Numerically, digits have a lower average pixel overlap than letters in Courier New (letters average: 35%; digits average: 29%) and Times New Roman (letters: 32%, digits: 29%), but not in Consolas (letters: 45%, digits: 47%) or Arial (letters: 37%, digits: 38%). The difference between letters and digits is only significant for Courier New (U=66.0, N_1=10, N_2=26, $p < .05$).

2.1.3.2. Overlap: feature. The feature overlap metric considered letter and digit parts, along with their positions within a letter-based frame, as described above. Average feature overlap of letters and digits were computed between each character and all other characters and then these values were compared between letters and digits. The feature descriptions are font-invariant, corresponding to the canonical form for each character, and therefore there is only one feature overlap metric rather than one for each font.

Averaging feature overlap across all letters gives an overlap of .13 (average overlap between any given letter and all other characters, range: 0–1) and .11 for digits (average overlap between any given digit and all other characters). The comparison between letters and digits failed to reach significance ($p > .1$).

2.1.4. Visual discriminability/closest competitor distance

Starrfelt and Behrmann (2011) developed a measure of discriminability for letters and digits which considers the closest competitors for a given character (i.e., other characters with pixel overlap over a set threshold). I considered a simplified notion of discriminability, defined by the overlap value of the closest competitor for each character. One advantage of this approach is that the number of competitors is not dependent on an overlap threshold. Instead, each character is evaluated by its similarity to its most similar character. However, the term discriminability is not quite appropriate in this context because higher values indicate a closer competitor, which would result in *harder* discrimination of a character rather than higher discriminability. Instead, this measure reflects the closest visual competitor, or simply competitor distance.

Comparison of digit and letter competitor distance involves two steps: For each character, I calculated the closest competitor character (the one with the highest overlap value), and then compared the values of the closest competitors for the letters to those for the digits. The pixel and feature overlap measures were both used, giving pixel competitor distance and feature competitor distance. These two measures diverge; for example, the highest pixel competitor for A (Consolas) considered out of all the characters is 4, with .65 pixel overlap. Considering features instead, the highest competitor for A (font independent) is a tie between V and X, both with .28 feature overlap.

2.1.4.1. Competitor distance: pixel. Considering pixel overlap across all characters, letters and digits do not differ in the value of their closest competitor in three of the fonts (*ps* > .1). However, they do differ in Times New Roman: Average competitor distance across letters is .68 and for digits is .61 (U=71.5, N_1=10, N_2=26, $p < .05$).

2.1.4.2. Competitor distance: feature. Letters and digits do not differ in competitor distance computed across all characters (U=116.5, N_1=10, N_2=26, $p > .6$). Letters had an average competitor distance of .44 and digits had an average competitor distance of .48.

Lower competitor distance for digits than letters was found with only a single font (Times New Roman) but not the other three fonts, nor the font-independent feature measure. When comparing the digits to letters using a measure that is not sensitive to an arbitrary threshold or the number of characters, distance to the closest visual competitor does not differ. Accordingly, this property of the characters – similarity to visual competitors – does not seem to underlie the digit advantage.

2.1.5. Visual complexity

Unlike visual overlap and visual competitor distance, visual complexity does not refer to relationships among characters (e.g., how easy is it to distinguish A from other characters) but to a property of a single character. It is another possible candidate for a visual property that might differ between digits and letters, and could thereby produce the

Table 1
Font-specific complexity comparisons.

	Number of pixels			Perimetric complexity		
Font	Digits	Letters	U	Digits	Letters	U
CONSOLAS 0123456789	1173.6	1224.8	111.5	63.4	68.4	109.0
ARIAL 0123456789	**1330.5**	**1608.6**	**70.5**	67.0	74.7	101.0
COURIER NEW 0123456789	**622.9**	**738.0**	**70.5**	**116.8**	**148.6**	**23.0**
TIMES NEW ROMAN 0123456789	**895.5**	**1251.9**	**44.0**	**76.2**	**107.2**	**22.0**

Significant comparisons ($p < .05$, 2-tailed, Mann-Whitney U test) are in bold.

digit advantage. Visual complexity was calculated three ways for each character: the number of black pixels (font-specific), the number of visual features, and perimetric complexity (font-specific). Complexity via all three methods was compared for the letter and digit sets.

2.1.5.1. Complexity: number of pixels. Table 1 details the number-of-pixels comparisons for each font. In Consolas, the number of pixels in letter and digit stimuli do not differ (*ps* > .1). Arial, Courier New and Times New Roman display significant differences in number of pixels between letters and digits (all *ps* < .05). In all three of these fonts, digits have fewer pixels than letters.

2.1.5.2. Complexity: number of features. The number of features for each character was counted from the feature set described above. The average number of features for letters is 2.65; the average number of features for digits is 2.3. These values do not differ significantly, U=97.0, N_1=10, N_2=26, $p > .2$.

2.1.5.3. Perimetric complexity. Perimetric complexity has been used by prior researchers as a metric of visual complexity of letter stimuli (Pelli et al., 2006; Ziegler et al., 2010). It is defined as the square of the perimeter of the ink area, divided by total ink area. This can be thought of as indexing the degree of intricacy of the lines comprising a character, relative to the character's size. The method described in Pelli et al. (2006) was implemented in MATLAB and computed from the same image files used for pixel overlap. In Consolas and Arial, the two sans-serif fonts, the perimetric complexity of letter and digit stimuli do not differ (*ps* > .1). However, Courier New and Times New Roman display significant differences in perimetric complexity between letters and digits (*ps* < .05). In both fonts, digits have lower perimetric complexity than letters; this is numerically true also for the sans-serif fonts (see Table 1).

2.1.6. Discussion

Visual similarity and discriminability comparisons were conducted between letters and digits to determine whether a difference in such a property might underlie the digit advantage. The necessary precondition for the visual property hypothesis is that digits must be less visually similar (confusable) or easier to discriminate than letters.

2.1.6.1. Pixel and feature overlap and discriminability. Pixel and feature overlap comparisons did not reveal stark differences between letters and digits. Pixel overlap measures vary widely across fonts and the comparisons between average pixel overlap of letters and digits were not statistically significant in three of four fonts (significant only in Courier New). These results suggest that there is no systematic difference between letters and digits in this type of visual similarity, and that it cannot explain the digit advantage.

The results of the feature-based visual similarity analyses are also straightforward: No differences were found between letters and digits in feature overlap. The feature set was designed to include features that are abstract, i.e., invariant to properties such as size, absolute position, and line thickness, and thus reflect proposed features at the character shape level. To the extent that this attempt was successful, a lack of difference between letters and digits on overlap at this level suggests that at the character shape level the set of digits and the set of letters differ only in terms of their connections to higher representations (i.e., allographs in the McCloskey and Schubert (2014) theory).

A version of discriminability, indexed by the strength of the closest competitor for a given character, was also calculated using pixel and feature similarity. Differences were only found between letters and digits using pixel overlap for Times New Roman, but not for any other fonts or the abstract feature set. These tests of the visual relationship between characters took into account the entire character set (letters + digits) to comply with the properties of the shared identification system. This is also in accordance with the observation that the digit advantage arises in both pure and mixed stimulus sets, when digits or letters cannot be pre-selected for activation due to context or expectations.

2.1.6.2. Visual complexity. Visual complexity was calculated in three ways: as number of pixels in each character, number of features in each character, and perimetric complexity. Considering features, the result is very straightforward: Letters and digits do not differ in the number of features that comprise them. For pixels and perimetric complexity, results varied by font. For both Courier New and Times New Roman, letters have higher average perimetric complexity and number of pixels than digits. These differences parallel the font types: Sans-serif fonts show no difference while serif fonts do (see Fig. 1 the characters in each font). In serif fonts, the serifs on the letters are more prominent and frequent than on digits (e.g., 6, 9 have no serifs in either font), which would reduce the number of pixels and perimetric complexity for digits relative to letters; this property of the fonts may explain the observed difference. In addition, Arial showed a significant difference between letters and digits in number of pixels, but not perimetric complexity.

Where digits and letters were found to differ, digits tend to have slightly numerically lower pixel overlap, pixel discriminability, number of pixels, and perimetric complexity than letters, in the direction required by the visual properties hypothesis. However, the lack of consistency across fonts is problematic for the hypothesis. The relative frequency of observing letters and digits in serif versus sans-serif fonts is not available, but to a reasonable approximation most text read on a screen uses sans-serif fonts while printed materials tend to use serif fonts. Furthermore, the studies reporting better digit than letter performance likely used a mix of fonts. Therefore, despite the difference found for the serif fonts, the lack of difference found with sans-serif fonts leaves us with no uniform reason for why digits are identified more rapidly and accurately than letters. The visual differences observed may be a general property of the visual forms of letters and digits in serif fonts, but visual differences are unlikely to be a primary driver of the digit advantage more generally.

2.2. Hypothesis 2: digits are more frequently encountered

The remaining hypotheses for the digit advantage refer to properties of letters and digits other than their visual appearance. Ingles and Eskes (2008) provided the hypothesis that differences in written frequency of occurrence for digits and letters might explain the performance discrepancy. Rath et al. (2015) also suggested that greater familiarity with single digits as opposed to single letters might explain higher identification accuracy. To provide an explanation for the superior performance on digits as a set, the average written frequency for digits would need to differ from that for letters. Presumably identification performance would be facilitated by a higher frequency of occurrence; if digits were more frequent than letters.

Effects of character frequency would suggest that the activation of allograph and/or abstract character identity representations is modulated by character frequency. This suggestion has already been made for letter identification, and letter frequency effects have been reported across a variety of single letter tasks (e.g., Jones and Mewhort, 2004; New and Grainger, 2011; Walker and Hinkley, 2003). However, digit frequency is not commonly investigated in this context. Two studies have investigated letter and digit frequency in the context of grapheme-color synaesthesia: Beeli et al. (2007) reported significant correlations between the saturation of synesthetic colors for each character and the character's frequency, and Smilek et al. (2007) reported significant correlations between luminance and character frequency. However, both of these studies considered the relative frequency of letter and digit sets separately; it is not known whether the relationship holds across the combined set of characters.

Another instance in which character familiarity has modulated performance is in the alphanumeric category effect in visual search (see Section 1.1.). Polk and Farah (1994) compared the performance of postal employees who sorted mail with postal codes containing digits only (e.g., 21218) to that of employees who sorted mail with codes containing both letters and digits (e.g., V6T 1Z4). They found an attenuated alphanumeric category effect for the employees with experience with mixed-category codes (Polk and Farah, 1994), suggesting that the identification system is sensitive to experience with the two character types. In summary, though evidence exists for effects of letter and digit frequency and co-occurrence of the character types, questions remain about the specific effect of familiarity on character identification.

2.2.1. Character frequency

Of published character frequency metrics (e.g., Benford, 1938; Jones and Mewhort, 2004; Mayzner and Tresselt, 1965; Solso and King, 1976), Jones and Mewhort (2004) is the only one to contain uppercase and lowercase letters as well as digits: These authors tallied digit and case-specific letter raw frequencies from the New York Times article archives (containing about 14 million words). In these comparisons I tested both cases separately as well as combined because it is unclear which letter case was used in much of the neuropsychological literature supporting the digit advantage (e.g., Holender and Peereman, 1987; Starrfelt and Behrmann, 2011).

Comparing lowercase letters to digits confirmed that on the whole, lowercase letters are more frequent than digits (Mann-Whitney $U=41.0$, $N_1=10$, $N_2=26$, $p<.01$). The average frequency of lowercase letters (2.37 million) exceeds that of digits (283,418) by nearly an order of magnitude. This result is inconsistent with the hypothesis that better performance for digits than lowercase letters could arise from digits' higher frequency. In the frequency ranking of lowercase letters and digits, all of the digits are ranked below 21 of the letters; there is some interleaving of the letter and digit frequency values at the lower end of the frequency ranks.

The discrepancy between the frequencies of uppercase letters and digits is much smaller and in the opposite direction: The average frequency of uppercase letters (135,677) is lower than that of digits (283,418). Comparing these sets indicates that digits are significantly more frequent than uppercase letters ($U=44.0$, $N_1=10$, $N_2=26$, $p<.01$). Unlike with lowercase letters, the four most frequent characters among the combined set of uppercase letters and digits are digits (0, 1, 5, 2), and half of the letters are less frequent than all of the digits.

Cross-case letter frequency (i.e., the average of lower- and uppercase frequencies) differs from both lowercase ($U=230.0$, $N_1=26$, $N_2=26$, $p<.05$) and uppercase frequency ($U=41.0$, $N_1=26$, $N_2=26$, $p<.001$). Cross-case letter frequency, like lowercase letter frequency, is significantly higher than digit frequency ($U=49.0$, $N_1=10$, $N_2=26$, $p<.01$).

2.2.2. Discussion

The analyses of character frequency revealed opposite patterns depending on letter case. On average, lowercase letters are more frequent than digits, while uppercase letters are less frequent than digits. Considering the two cases combined, letters are more frequent than digits. Across both cases there is a large amount of intermixing of character frequencies such that some digits are more frequent than some letters and some digits are less frequent than some letters. The frequency hypothesis depends on digits having higher frequency than letters, which is true exclusively for uppercase letters, suggesting that this hypothesis is not viable. Overall, the lack of systematic difference in character frequencies is inconsistent with frequency as the cause of the digit advantage.

2.3. Hypothesis 3: fewer digit identities boost identification

The third hypothesis for the superiority of digit identification contends that the smaller set of digit identities facilitates the process of activating the target digit. A few authors have suggested this possibility, though generally without describing the mechanism of facilitation (Cohen and Dehaene, 1995; Ingles and Eskes, 2008; Piazza and Eger, 2015; Polk et al., 2002; Polk and Farah, 1998). Starrfelt and Behrmann (2011) made this explanation more explicit, referring to “guessing rate” (p. 2292), proposing that the digit advantage may result from the higher chance rate of correctly guessing one of ten possible digits rather than one of twenty-six possible letters. However, the guessing rate explanation presupposes that the identification system contains knowledge that the target is a digit and accordingly no letter identities are being considered (or vice-versa). This explanation is not compatible with evidence suggesting that character identification is not constrained by knowledge of the category of the stimulus (McCloskey and Schubert, 2014). Given this property of the shared identification system, the smaller number of digit than letter identities seems unlikely to contribute to the digit advantage.

Another possible interpretation of the guessing hypothesis is that higher-level knowledge, such as knowing that the current block of experimental trials you are completing only contains digits, can affect the outcome of the identification process. This might happen via a monitoring process after identification (but prior to production of the response), or by feedback into the identification system from elsewhere in the cognitive system. Such a mechanism likely operates in certain contexts, but it is not obvious why this would preferentially boost digit performance over letter performance. Additional details would be needed to explain how this type of information would interact with the number of character identities to produce an advantage for situations when digits are predicted than when letters are predicted.

In summary, the hypothesis that the smaller number of digits provides a basis for the digit advantage seems unlikely. A more explicit and testable version of the hypothesis would need to contend with existing data suggesting that the digit advantage persists in mixed-category contexts and that letter/digit category knowledge does not appear to constrain activation of character identities within the shared identification system.

2.4. Hypothesis 4: digits are supported by semantics, and/or the right hemisphere

The final hypothesis that has been proposed to explain the digit advantage refers to connections between letters, digits and representations of meaning. Despite their visual similarities, digits and letters differ in their semantic content. Single digits are logographic symbols that convey a meaning, while letters typically convey meaning only when combined (single-letter words excepted) (Deloche and Seron, 1987; Holender and Peereman, 1987). For example, '7' indicates a semantic quantity, while 'T' does not have any semantic value.

Related to this property of the characters is the fact that when children learn digit forms they are mapping a visual symbol onto their pre-existing knowledge of a number, including the number word and representation of quantity (Benoit et al., 2013; Hurst et al., 2016). Digit forms are learned largely after children have acquired the ability to count and understand magnitudes (Benoit et al., 2013; Berteletti et al., 2010; Hurst et al., 2016; Sinclair et al., 1983). This background may favor fast acquisition of digit forms. By contrast, letter forms are not mapped onto existing knowledge, but rather are learned along with acquisition of the alphabetic principle (correspondence between written letters and the sounds of language) and the development of phonemic awareness in spoken tasks (Castles and Coltheart, 2004; Castles et al., 2009). The impact of early learning and the strong relationship between digits and semantic content could conceivably result in a digit identification advantage later in life.

Based on these considerations, some authors have suggested that there might be a more direct, or stronger, connection between digit identification and conceptual processing than between letter identification and conceptual processing (Cohen and Dehaene, 1995; Ingles and Eskes, 2008; Rath et al., 2015; Starrfelt and Behrmann, 2011; for a related proposal see: Miozzo and Caramazza, 1998). Therefore, these authors contend, there may be stronger feedback supporting digit than letter identification, leading to more accurate or faster identification performance.

Top-down effects on identification have also been proposed to impact letter identification, facilitating identification in certain situations. One prominent example is the word superiority effect, which refers to higher accuracy for identifying a letter when it is presented within a word relative to when it is presented as a single character (Reicher, 1969; Wheeler, 1970). This effect has been explained by appealing to interactivity between word and letter levels boosting letter identification within a word context, with no such word-level information available in single letter identification (e.g., McClelland and Rumelhart, 1981). Similar effects may be at work in digit identification, with an identification advantage produced by top-down effects from semantic/quantity representations.

On this point, it is relevant to consider whether numerical quantity representations are automatically activated by visually-presented digits. Some researchers have suggested that, unlike in arithmetic or magnitude comparison tasks, digit naming may not require access to a semantic (quantity) representation (e.g., Dehaene, 1992; Deloche and Seron, 1987). The automaticity of semantic access has been assessed using tasks which do not require this access (e.g., judging whether two digits are physically identical, or whether a digit is identical to the digit '5'), and access to quantity would facilitate or interfere with performance. Researchers have argued against automatic semantic activation (e.g., Cohen, 2009; Wong and Szücs, 2013), in favor of it (e.g., Ganor-Stern and Tzelgov, 2008), or for automatic but slow access (García-Orza et al., 2012). Accordingly, it is unclear whether there is automatic access to semantics for digits in all situations, and if so, what the indirect effect on identification performance might be.

One test of the relevance of semantic feedback might be possible with an individual who has damage to higher level knowledge for digits but not letters, perhaps as an isolated deficit to numerical quantity knowledge. This hypothetical individual could then be tested on the relative speed and accuracy of letter and digit identification. To my knowledge, this specific alignment of impairment and research question have not occurred, but it is an open question for future work. However, some versions of the digit semantics hypothesis, which refer to the neural instantiation of semantic and identification processes, can be considered in the light of data from neuroimaging studies.

2.4.1. Neural bases of number semantics

One version of the hypothesis that digit identification performance is boosted by semantics further specifies that the boost is due to the neural instantiation of digit semantics in the right hemisphere or in both hemispheres, contrasted with a unilateral left representation for letters. In a review of the neuropsychological literature on letter and digit naming, Holender and Peereman (1987) suggest that the right hemisphere stores some semantic representations for digits, while letter and word semantics are stored in the left hemisphere. They posit that the right hemisphere may provide support for digit naming when the left hemisphere is damaged, leading to a digit advantage in these individuals (see also: Deloche and Seron, 1987). Cohen and Dehaene (1995) made a similar suggestion, positing that digits have bilateral representations (of visual form and magnitude) while letters are represented mainly in the left hemisphere. Thus, they contend, in the case of unilateral brain damage, digits will tend to be fully or more preserved.

The representation of number semantics (i.e., numerical quantity representation) is generally bilateral and involves a large network of parietal, frontal, and cingulate regions, with the degree of laterality varying among individuals and by task (Butterworth, 2005; Chochon and Inserm, 1999; Park et al., 2011; Piazza and Eger, 2015; Prado et al., 2011). This literature suggests that different numerical abilities may be subserved by different brain areas, but importantly that both hemispheres are involved. By contrast, it is generally agreed that linguistic processing, including for reading, is mainly limited to the left hemisphere (Martin et al., 2015; Taylor et al., 2013; but see: Hickok (2013), Hickok and Poeppel (2004) for evidence of bilaterality in speech perception). It is possible that bilateral representation of number semantics would result in stronger feedback connections to digit identification than those coming from unilateral letter/word processing, though a more thorough understanding of neural connectivity and communication would be needed to fully evaluate this particular hypothesis.

2.4.2. Neural bases of digit identification

An alternative formulation of this hypothesis is that digit *identification* specifically is localised in the right hemisphere and therefore tends to be preserved in individuals with left brain damage. Naturally, this explanation cannot account for the presence of digit identification deficits in individuals with unilateral left hemisphere lesions. Furthermore, it does not predict the lack of double dissociation between digit and letter identification, because anatomical independence would allow each system to be damaged independently. And finally, turning to unimpaired individuals, this hypothesis cannot account for the digit advantage there without further assumptions about the impact of hemispheric localization on normal performance.

2.4.2.1. Co-occurrence hypothesis. A related hypothesis is the co-occurrence hypothesis, developed by Polk and Farah (1995, 1998) to explain the neural localization of identification processes. They posit that the contexts in which letters and digits are encountered result in distinct brain areas for letter and digit identification. Specifically, letters are often identified in close in space and time to other letters (e.g., in books), but less often close to digits. Polk and Farah (1995, 1998) hypothesise that the result of this environmental segregation is neural segregation, brought about by Hebbian learning. Neural segregation could explain how letter identification deficits can occur absent a concomitant digit identification deficit, and why letters are often more impaired when the deficits do co-occur. Results of a neural network simulation support the assertion that when inputs are segregated by character type distinct sub-areas arise for the representation of the two types (Polk and Farah, 1995). However, the simulation is simplified in a number of ways which limit the implications for neural representation.[4] Furthermore, it is difficult to assess the degree of letter and digit segregation in the environment. However, recent neuroimaging studies localizing brain regions for letter and digit processing can be brought to bear on the question of neural segregation.

2.4.2.2. Evidence from neuroimaging. Letter (and word) identification has been localised (using fMRI) mainly to the left hemisphere, including the so-called Visual Letter Area (e.g., Rothlein and Rapp, 2014; Thesen et al., 2012). Digit identification substrates have been researched relatively more recently, and early fMRI studies often found no neural preference for digit over letter stimuli, or a mix of left- and right-hemisphere preference (Allison et al., 1994; James et al., 2005; Polk et al., 2002; Polk and Farah, 1998; Price and Ansari, 2011; Reinke et al., 2008; Roux et al., 2008).

Some researchers have found regions of the right ventral visual stream showing a preferential response to single digits over letters (Cantlon et al., 2011; James et al., 2005). Shum et al. (2013) tested the responses of intracranial electrodes to single digit, single letter, pseudoletter, pseudodigit, and word stimuli. They discovered five right-hemisphere electrode sites that showed a preference for digits. Additional sites that responded more to digits than other characters (letters and pseudo-characters) were found in the inferior temporal/fusiform region of both hemispheres (Shum et al., 2013). Following up on these results, Abboud and colleagues (Abboud et al., 2015) reported functional connectivity analyses of partially-distinct processing networks for letters and digits, mainly involving the left and right hemispheres, respectively, including a right hemisphere digit-preferential Visual Number Form Area (VNFA). However, this finding was not replicated in a follow-up study by another group (Peters et al., 2015).

A subsequent study found bilateral preference for digits over other visual stimuli; this preference was stronger in the right hemisphere when digits were compared to pseudodigits, but stronger in the left hemisphere when digits were compared to other visual stimuli (Grotheer et al., 2016b). Grotheer and colleagues (Grotheer et al., 2016a) applied TMS to the VNFA during a familiar character (letters and digits) vs. unfamiliar character (pseudoletters and pseudodigits) decision task. TMS to the right VNFA reduced task performance, while TMS to the vertex and to a left hemisphere object-preferential region did not (Grotheer et al., 2016a). The authors concluded that the right VNFA was causally involved in digit recognition, but importantly, their results do not suggest that the region is selective to digits over letters.[5] The presence of a right-hemisphere region which is involved in both letter and digit processing but more strongly in digit processing is also supported by a study of split-brain patients, which revealed higher accuracy on digit than letter identification when presented to the right hemisphere (Teng and Sperry, 1973). However, in this study both character types were identified with some success, suggesting that the right hemisphere is not exclusive to digit processing.

Taken together, these results suggest that though digit identification may predominate in the right hemisphere, right hemisphere areas involved in digit identification are not selective to digits. The most stringent test of the role of the VNFA (Grotheer et al., 2016a) revealed that it was causally involved in letter recognition as well (for relevant discussion see: Merkley et al., 2016). Thus, current evidence is consistent with bilateral identification processing of digits and letters, suggesting that the precondition for this version of the right hemisphere hypothesis is not true. Though it remains possible that the right hemisphere makes a stronger contribution to digit identification than letter identification, this hypothesis would also require additional explication to be fully tested.

2.4.3. Discussion

Suggestions that digit identification might be boosted by stronger connections to number semantics, or that the digit representations in the right hemisphere improve performance do not seem to straightforwardly map onto the digit advantage. The existence of more semantic representation for digits and letters may be accurate, but its impact is difficult to evaluate. In terms of neural instantiation, number semantics and digit identification processes have been localised to both hemispheres. Letter identification is mainly confined to the left hemisphere; the combination of these facts may explain why cases of digit identification impairment without letter identification impairment have not been reported. However, as an explanation for higher accuracy or faster responding among unimpaired individuals, these hypotheses would need to be extended and the impact of particular neural bases on performance clarified.

3. General discussion

Four main hypotheses have been proposed to explain the observed discrepancy between digit and letter identification. However, none of the hypotheses has provided a satisfactory answer. Some hypotheses require unsupported assumptions, such as that the category of a character constrains the identification processes. Others are not backed up by data to explain the phenomenon. A number of the results in this study provided partial or suggestive evidence for a role of visual properties or character frequency in facilitating digit over letter identification, but overall a consistent picture did not emerge.

The first four variables considered were visual overlap, visual discriminability, visual complexity, and frequency of occurrence. The first three were computed for both pixel- and feature-based metrics, which are assumed to index different levels of processing in the identification system. I found that digits in serifed fonts tended to have lower pixel overlap and complexity than letters in those fonts, but no differences were found in terms of features. As discussed previously, the digit advantage has been reported across a range of fonts, and the underlying cause of the advantage should therefore be font-independent.

In character frequency, digits are less frequent than lowercase letters but more frequent than uppercase letters. This is a straightforward consequence of prose containing mainly lowercase letters (in English). Studies reporting a digit advantage have used a mix of uppercase and lowercase letters, suggesting that even the significant difference between uppercase letter and digit frequency cannot explain the digit advantage. Instead, the underlying cause should be consistent across letter case.

Having determined that digits and letters do not systematically differ in visual properties or frequency (in the required direction), we are left with set size and support from higher level processing or the right hemisphere as possible contributors to the digit identification advantage. The hypothesis that the digit advantage is caused by the smaller number of digit than letter identities is difficult to assess empirically. However, this hypothesis is inconsistent with theories of a shared alphanumeric identification system in which characters are identified without regard to their category. This lack of category-

[4] The limitations of the simulation for neural letter and digit processing are three-fold: the input representations do not capture visual similarity (or any other similarity) between characters, equal numbers of "letter" and "digit" inputs were employed, and each input was presented with equal frequency. These aspects run counter to the properties of the characters and the statistics of occurrence in the environment, and the biasing effect of their combined influence is unclear.

[5] Grotheer et al. (2016a) were puzzled by the impact of rVNFA TMS on letter recognition. However, according to the shared identification system account proposed by McCloskey and Schubert (2014), the task they used relies on the allograph level of representation (prior to activation of letter and digit abstract identities), which does not distinguish between letter and digit processing. Accordingly, the finding of an area that is causally involved in this task for both letters and digits is wholly consistent with McCloskey and Schubert's proposal for the character identification system.

restriction is a key principle of the shared system proposed by McCloskey and Schubert (2014), and is also consistent with Norris et al.'s (2010) and Grainger et al. (2016a, 2016b) models. Thus, none of the current shared system proposals seem to permit an explanation of the digit advantage on the basis of the number of letter and digit identities. There may be some contexts in which the identification system is restricted to letter or digit identities, but this exception certainly does not apply universally.

The final hypothesis considered concerns the impact of semantic knowledge on identification and the relationship between cognitive processing and neural substrates. Because single digits have semantic content while single letters do not (with the possible exception of A and I), the suggestion is that feedback from semantics would boost digit over letter identification. In the case of the digit advantage after brain damage, a specific version of this hypothesis states that right hemisphere areas support digit semantics and/or identification but not letter processing. Digit semantics seems to be instantiated bilaterally, and it is unclear to what extent semantic information for digits impacts on identification processes in impaired or unimpaired individuals. In identification, there is a growing body of evidence in support of a right-hemisphere region that responds preferentially to digits over letters. However, the most stringent test of this region's processing suggests that it is causally involved in both digit and letter identification. In summary, the role of semantics in digit identification remains unclear, as does the role of right hemisphere processing. Future research will certainly help clarify the neural bases of numerical processing and neural overlap between letter and digit identification, which may provide additional relevant evidence for the right hemisphere/feedback hypotheses.

3.1. What is the digit advantage?

Critically, all four hypotheses were proposed to explain a difference in identification performance across the entire set of digit and letter characters. However, considering set-wise accuracy or RT may obscure differences in accuracy across individual characters. The phenomenon of higher digit than letter accuracy may only be true for the mean accuracy, not for all digits compared to all letters; this level of performance detail has generally not been reported. For example, the acquired dyslexic individual reported by McCloskey and Schubert (2014) showed average identification accuracy for digits that exceeded her average accuracy for letters (a digit advantage). However, there was considerable overlap in the accuracy for individual digit and letter characters. Due to a lack of similar data for other individuals (both impaired and unimpaired), we are currently in the situation of attempting to explain a digit advantage which may be more accurately described as an advantage for particular characters, many of which happen to be digits. Until this data has been provided, attempts to explain the digit advantage will always be at a disadvantage.

It may be that the digit advantage is better understood not as an advantage for one category of stimuli over another, but faster or more robust processing for particular characters, due to particular factors relevant in the identification system. As discussed in the introduction (Section 1.2.), it is well known that letters are not all identified with the same speed and accuracy. This type of differentiation across stimuli of the same type is also found in other domains: In lexical processing (e.g., lexical decision) particular words are processed more quickly than others, due to a variety of reasons (frequency, neighbourhood size, regularity, etc.). It seems then that the observation that digits show an identification advantage may be no different than the observation that DOG is faster to read than VOW: a result of properties of these items (frequency, in this example) rather than a categorical difference in processing.

3.2. Predicting performance on an individual character level

Detailed investigation of letter and digit identification depends on authors of future studies to report and analyze dependent measures for individual digit and letter characters. This will allow us to determine whether the conceptualization of the digit advantage is correct, or whether it is a generalization that is only true on aggregate. At that point, more specific hypotheses can be considered for the observation that digits tend to out-perform letters, and more sophisticated analyses can be conducted to evaluate the hypotheses. Hypotheses which currently appear weak may gain more explanatory power when considered on the level of individual characters. Accordingly, stronger conclusions about the source of the digit advantage could be drawn by comparing by-character performance to by-character measures of frequency, visual similarity, and complexity (which were collapsed across sets here). Alternatively, none of these hypotheses may be able to account for the digit advantage alone, but may in combination. For example, a small effect of visual similarity may combine with a small effect of visual complexity to result in higher accuracy for 1 than W. These combinations of factors should be tested against a dependent measure of accuracy or RT for individual characters, rather than the average of two large sets.

In line with the suggestion for more in-depth consideration of letter and digit identification and a renewed focus on character-specific performance, the by-character values for all of the visual characteristics considered in this paper (e.g., average pixel overlap, perimetric complexity) are presented in Appendix B. The purpose of including these data is twofold. First, it allows researchers to examine the relative ordering of letters and digits on each metric and observe the degree to which the two categories are intermixed. Second, it allows for the construction of stimulus sets which are matched on any of these variables, should the need arise in future experiments comparing letters and digits.

3.3. Theoretical considerations

The ability of particular variables to predict identification accuracy and response times might also inform models of the functioning of the cognitive character identification system. Though three competing models of alphanumeric character identification have been proposed, none of these have been described or instantiated in sufficient detail to permit an understanding of the effects of variables such as frequency and visual similarity. Furthermore, while theories of numerical cognition and magnitude representation exist, they are largely divorced from consideration of the early stages of digit identification, and vice-versa. A fully-unified theory of letter and digit processing, from early visual features to semantic/lexical representations could help predict and understand the effects of variables on identification performance. In support of this goal, identification performance also helps to constrain models of character identification (for recent examples in the context of children's acquisition of letter and digit processing, see: Grainger et al., 2016a; Schubert et al., submitted). With additional details that establish the effects of properties of context, visual properties, and top-down influences, identification theories such as those discussed in the introduction can be modified and tested for their compatibility with the performance advantage of certain characters over others.

3.4. A note on other digit notations

This paper has focused on the comparison between Arabic digits and Roman letters, reflecting the preponderance of research and theory-building based on these two notational systems. However, research has also been conducted using other numerical notations – such as the Arabic-Indian numerals and Persian-Indian numerals – to investigate issues such as automatic access to numerical quantity (Section 2.4.). The extent to which a digit advantage exists in other orthographies (e.g., Arabic-Indian numerals compared to Arabic script) is currently unclear. Hypotheses for the digit advantage which depend on visual and frequency properties of the characters obviously would need to be re-evaluated for a new orthography, while those depending on semantic and neural properties may index more universal properties of digit processing. However, by-character analyses are likely to produce the most effective means to evaluate character identification across orthographies.

3.5. Conclusion

This study has revealed that the hypotheses that have been proposed to explain the digit advantage generally fall short of that aim. However, promising avenues exist to explore the hypotheses as they apply to performance on individual characters. The by-character approach promises not only to explain a puzzling discrepancy between letter and digit performance, but to constrain and enrich theories of alphanumeric character identification more broadly.

Funding

This work was supported by the Australian Research Council Centre of Excellence for Cognition and its Disorders (CE110001021), www.ccd.edu.au.

Acknowledgements

Thanks are due to M. McCloskey and R. Starrfelt, as well as two anonymous reviewers, for their helpful feedback on this work.

Appendix A. Character feature set

See Tables A1–A3.

The feature classes used are Orthogonal, Slant, and Curve. These describe the large categories of visual features found in uppercase letters. Within the Orthogonal class are Horizontal and Vertical features; within the Slant class are Slant-left and Slant-right features, and within the Curve class are Curve-facing-up, Curve-facing-down, Curve-facing-left, Curve-facing-right, and Curve-Closed (loop) features. Each letter is considered within a rectangle in letter-space, which has a vertical height of Full and a horizontal width of Half (Half is also used as the size specification for a feature which extends half the height in the vertical dimension). Slant lengths are approximate; a slant extending from the top to bottom of letter-

Table A1
Features for uppercase letters and digits.

Character	Feature class	Feature	Size	Position 1	Position 2
A	Slant	Slant left	Full	CM	RB
	Slant	Slant right	Full	CM	LB
	Orthogonal	Horizontal	Half	LM	RM
B	Orthogonal	Vertical	Full	LT	LB
	Curve	Curve facing left	180	LT	LM
	Curve	Curve facing left	180	LM	LB
C	Curve	Curve facing right	180	RT	RB
D	Orthogonal	Vertical	Full	LT	LB
	Curve	Curve facing left	180	LT	LB
E	Orthogonal	Vertical	Full	LT	LB
	Orthogonal	Horizontal	Half	LT	RT
	Orthogonal	Horizontal	Half	LM	RM
	Orthogonal	Horizontal	Half	LB	RB
F	Orthogonal	Vertical	Full	LT	LB
	Orthogonal	Horizontal	Half	LT	RT
	Orthogonal	Horizontal	Half	LM	RM
G	Orthogonal	Vertical	Half	RM	RB
	Orthogonal	Horizontal	Half	CM	RM
	Curve	Curve facing right	180	RT	RB
H	Orthogonal	Vertical	Full	LT	LB
	Orthogonal	Vertical	Full	RT	RB
	Orthogonal	Horizontal	Half	LM	RM
I	Orthogonal	Vertical	Full	CT	CB
	Orthogonal	Horizontal	Half	LT	RT
	Orthogonal	Horizontal	Half	LB	RB
J	Orthogonal	Vertical	Full	RT	RB

(continued on next page)

Table A1 *(continued)*

Character	Feature class	Feature	Size	Position 1	Position 2
	Orthogonal	Horizontal	Half	LT	RT
	Curve	Curve facing up	180	LM	RM
K	Orthogonal	Vertical	Full	LT	LB
	Slant	Slant left	Half	RT	LM
	Slant	Slant right	Half	LM	RB
L	Orthogonal	Vertical	Full	LT	LB
	Orthogonal	Horizontal	Half	LB	RB
M	Slant	Slant left	Full	LT	LB
	Slant	Slant left	Half	LT	CM
	Slant	Slant right	Half	RT	CM
	Slant	Slant right	Full	RT	RB
N	Orthogonal	Vertical	Full	LT	LB
	Orthogonal	Vertical	Full	RT	RB
	Slant	Slant left	Full	LT	RB
O	Curve	Curve closed	360	CT	CB
P	Orthogonal	Vertical	Full	LT	LB
	Curve	Curve facing left	180	LT	LM
Q	Curve	Curve closed	360	CT	CB
	Curve	Curve right	90	CB	CR
R	Orthogonal	Vertical	Full	LT	LB
	Curve	Curve facing left	180	LT	LM
	Slant	Slant left	Half	LM	RB
S	Curve	Curve facing left	180	RT	CM
	Curve	Curve facing right	180	CM	LB
T	Orthogonal	Vertical	Full	CT	CB
	Orthogonal	Horizontal	Half	LT	RT
U	Orthogonal	Vertical	Full	LT	LB
	Orthogonal	Vertical	Full	RT	RB
	Curve	Curve facing up	180	LB	RB
V	Slant	Slant left	Full	LT	CB
	Slant	Slant right	Full	RT	CB
W	Slant	Slant left	Full	LT	LB
	Slant	Slant left	Half	CM	RB
	Slant	Slant right	Half	CM	LB
	Slant	Slant right	Full	RT	RB
X	Slant	Slant left	Full	LT	RB
	Slant	Slant right	Full	RT	LB
Y	Orthogonal	Vertical	Half	CM	CB
	Slant	Slant left	Half	LT	CM
	Slant	Slant right	Half	RT	CM
Z	Orthogonal	Horizontal	Half	LT	RT
	Orthogonal	Horizontal	Half	LB	RB
	Slant	Slant right	Full	RT	LB
0	Curve	Curve closed	360	CT	CB
1	Slant	Slant right	Half	LM	RT
	Orthogonal	Vertical	Full	CT	CB
	Orthogonal	Horizontal	Half	LB	RB
2	Slant	Slant right	Full	RT	LB
	Orthogonal	Horizontal	Half	LB	RB
	Curve	Curve down	180	LM	RM

(continued on next page)

Table A1 *(continued)*

Character	Feature class	Feature	Size	Position 1	Position 2
3	Curve	Curve left	180	LT	LM
	Curve	Curve left	180	LM	LB
4	Slant	Slant right	Full	CT	LM
	Orthogonal	Vertical	Full	RT	RB
	Orthogonal	Horizontal	Half	LM	RM
5	Orthogonal	Vertical	Half	LT	LM
	Orthogonal	Horizontal	Half	LT	RT
	Curve	Curve left	180	LM	LB
6	Curve	Curve closed	360	CM	CM
	Curve	Curve right	90	LM	RT
7	Orthogonal	Horizontal	Half	LT	RT
	Slant	Slant right	Full	RT	LB
8	Curve	Curve closed	360	CT	CT
	Curve	Curve closed	360	CM	CM
9	Curve	Curve closed	360	CT	CT
	Curve	Curve left	90	RM	LB

Table A2
Rules for computing feature overlap.

Overlap type	Units of overlap assigned
Feature class only	.5
Feature class and feature	1
Feature class, feature, and feature size	2
Feature class, feature, feature size, and position	3 (maximum)

space gets a length of Full even though it is physically longer than an Orthogonal feature extending from the top to bottom. Curves have a size parameter indicating the length in degrees covered by the curve.

To specify positions in letter-space, the FullxHalf grid is subdivided into 4 sections, with the vertices of the grid labeled by their position in the horizontal dimension (Top, Middle, Bottom) and in the vertical dimension (Left, Center, Right). These vertices are abbreviated by their first letter, and each feature is defined by two positions approximating the two ends of the feature. The system is arbitrarily set such that Position 2 is always downward and/or to the right of Position 1. For example, the Slant left feature in A is described by Position 1 of Center-Middle (CM) and Position 2 of Right-Bottom (RB).

To generate the overlap matrix for all characters, the rules in Table A.2 were followed for each feature in every pairwise combination of characters. The overlap values for a character-pair where then summed and divided by the following: 3 times the product of the number of features in each of the two characters. This produces overlap values in the range 0–1, with 1 overlap for identical characters.

Table A3
Feature overlap matrix (lower triangle only).

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0																																				
1	0																																			
2	.06	.17																																		
3	.17	0	0																																	
4	0	.26	.17	0																																
5	.06	.15	.11	.19	.15																															
6	.17	0	.06	.17	0	.06																														
7	0	.19	.28	0	.25	.19	0																													
8	.33	0	.06	.17	0	.06	.50	0																												
9	.17	0	.06	.25	0	.08	.29	0	.5																											
A	0	.15	.17	0	.22	.09	0	.17	0	0																										
B	.11	.09	.06	.56	.09	.24	.11	.03	.11	.17	.02																									
C	.17	0	.06	.17	0	.06	.25	0	.17	.17	0	0																								
D	.08	.14	.06	.33	.14	.19	.08	.04	.08	.13	.03	.39	.08																							
E	0	.31	.21	0	.31	.28	0	.31	0	0	.21	.13	0	.19																						
F	0	.28	.17	0	.31	.28	0	.31	0	0	.20	.15	0	.22	.47																					
G	.06	.15	.11	.06	.15	.20	.08	.14	.06	.06	.07	.09	.33	.11	.13	.13																				
H	0	.28	.11	0	.35	.20	0	.17	0	0	.15	.2	0	.31	.43	.46	.20																			
I	0	.35	.2	0	.28	.28	0	.31	0	0	.17	.11	0	.17	.51	.48	.24	.39																		
J	.06	.19	.11	.06	.22	.2	.06	.19	.06	.06	.09	.13	.06	.17	.31	.31	.17	.31	.31																	
K	0	.19	.11	0	.19	.06	0	.08	0	0	.17	.11	0	.17	.13	.15	.06	.2	.11	.09																
L	0	.33	.19	0	.28	.22	0	.21	0	0	.14	.19	0	.29	.50	.47	.22	.47	.47	.28	.19															
M	0	.11	.14	0	.11	0	0	.17	0	0	.22	0	0	0	0	0	0	0	0	0	.22	0														
N	0	.20	.06	0	.24	.11	0	.08	0	0	.13	.19	0	.28	.22	.26	.11	.41	.22	.22	.24	.28	.11													
O	1	0	.06	.17	0	.06	.17	0	.33	.17	0	.11	.17	.08	0	0	.06	0	0	.06	0	0	0	0												
P	.08	.14	.08	.42	.14	.19	.08	.04	.08	.13	.03	.44	.08	.42	.19	.22	.08	.31	.17	.17	.17	.25	0	.28	.08											
Q	.58	0	.06	.17	0	.06	.29	0	.25	.17	0	.11	.17	.08	0	0	.08	0	0	.06	0	0	0	0	.58	.08										
R	.06	.11	.06	.28	.11	.13	.06	.06	.06	.08	.07	.30	.06	.28	.13	.15	.11	.20	.11	.11	.20	.17	.10	.22	.06	.33	.06									
S	.17	0	.06	.33	0	.14	.21	0	.17	.21	0	.28	.42	.21	0	0	.14	0	0	.06	0	0	0	0	.17	.08	.17	.14								
T	0	.33	.14	0	.28	.28	0	.29	0	0	.14	.14	0	.21	.46	.47	.22	.42	.53	.33	.14	.42	0	.28	0	.21	0	.14	0							
U	.06	.19	.06	.06	.22	.13	.06	.06	.06	.06	.04	.22	.06	.31	.22	.26	.13	.41	.22	.30	.19	.33	0	.37	.06	.31	.06	.20	.06	.28						
V	0	.08	.14	0	.14	0	0	.21	0	0	.28	0	0	0	0	0	0	0	0	0	.17	0	.33	.14	0	0	0	.08	0	0	0					
W	0	.11	.11	0	.11	0	0	.17	0	0	.22	0	0	0	0	0	0	0	0	0	.22	0	.33	.11	0	0	0	.11	0	0	0	.33				
X	0	.08	.19	0	.14	0	0	.29	0	0	.28	0	0	0	0	0	0	0	0	0	.17	0	.33	.19	0	0	0	.08	0	0	0	.42	.33			
Y	0	.15	.07	0	.11	.09	0	.11	0	0	.13	.02	0	.06	.07	.07	.09	.09	.07	.06	.22	.06	.28	.13	0	.11	0	.13	0	.08	.07	.17	.22	.17		
Z	0	.26	.30	0	.22	.22	0	.44	0	0	.24	.04	0	.06	.42	.37	.19	.22	.41	.22	.09	.33	.1	.09	0	.06	0	.06	0	.33	.07	.14	.11	.19	.09	

Appendix B. By-character metrics

see Tables B1–B5.

Table B1
Font-dependent visual properties: Arial.

Character	Average visual overlap: Pixel	Competitor distance: Pixel	Number of pixels	Perimetric complexity
0	.4215	.8119	1431	70.223
1	.2053	.5293	801	43.812
2	.4057	.6511	1323	68.178
3	.4289	.847	1288	69.876
4	.2827	.5196	1297	62.919
5	.4292	.7961	1437	77.476
6	.4513	.8096	1565	76.791
7	.3112	.5989	1004	52.384
8	.4573	.847	1607	71.935
9	.4315	.8119	1552	76.395
A	.3137	.5341	1555	69.608
B	.512	.7997	2099	83.773
C	.4334	.7937	1505	68.04
D	.4501	.7567	1826	82.586
E	.4778	.7821	1706	98.374
F	.4168	.7422	1358	66.716
G	.4219	.8314	1849	88.856
H	.3728	.7638	1700	83.015
I	.2045	.729	718	36.552
J	.3086	.6164	938	50.665
K	.3798	.639	1701	62.991
L	.3332	.5296	998	53.314
M	.269	.5722	2448	118.38
N	.3698	.7638	1992	79.387
O	.4043	.9065	1786	73.373
P	.4274	.7685	1563	70.097
Q	.3997	.9065	2002	82.877
R	.4407	.7685	1945	83.087
S	.4867	.7997	1710	76.634
T	.3035	.729	1079	59.949
U	.3944	.7567	1565	74.592
V	.2869	.5722	1278	68.249
W	.2638	.5341	2321	128.13
X	.3249	.5786	1507	59.192
Y	.2853	.492	1120	51.286
Z	.42	.6534	1554	72.073

Table B2
Font-dependent visual properties: Consolas.

Character	Average visual overlap: Pixel	Competitor distance: Pixel	Number of pixels	Perimetric complexity
0	.5432	.8234	1534	67.032
1	.386	.7396	977	57.007
2	.4621	.6794	1071	55.741
3	.4711	.7427	1024	65.846
4	.3504	.6507	1194	65.193
5	.5027	.8156	1087	70.248
6	.5061	.7654	1260	67.979
7	.3859	.6137	855	52.732
8	.5565	.8356	1489	62.611
9	.4992	.6921	1245	69.584
A	.3562	.6507	1314	63.709
B	.5824	.8356	1491	75.568
C	.4631	.8041	996	55.92
D	.5195	.8001	1428	70.519
E	.527	.8695	1148	79.973
F	.4717	.8695	998	60.637
G	.5128	.8041	1327	76.365
H	.4539	.7715	1367	76.635
I	.4148	.8559	1024	65.34
J	.4264	.726	921	56.278
K	.4289	.6183	1221	64.516
L	.3809	.7052	772	52.159

(continued on next page)

Table B2 *(continued)*

Character	Average visual overlap: Pixel	Competitor distance: Pixel	Number of pixels	Perimetric complexity
M	.406	.6978	1542	99.483
N	.4699	.7603	1516	86.282
O	.5037	.9048	1354	63.838
P	.4845	.7467	1124	63.9
Q	.4593	.9048	1630	76.303
R	.5265	.7843	1383	68.741
S	.5075	.8156	1137	63.17
T	.3449	.8559	878	53.797
U	.4699	.7853	1246	70.318
V	.3477	.5119	1151	61.474
W	.4048	.7014	1518	99.855
X	.384	.6184	1228	57.33
Y	.3419	.5315	955	51.297
Z	.4781	.6794	1175	64.831

Table B3
Font-dependent visual properties: Courier New.

Character	Average visual overlap: Pixel	Competitor distance: Pixel	Number of pixels	Perimetric complexity
0	.3288	.6349	671	123.33
1	.2983	.8593	447	110.26
2	.3499	.566	617	111.25
3	.3306	.6691	588	108.29
4	.227	.4068	684	118.47
5	.3035	.5693	638	125.53
6	.2582	.4394	691	121.43
7	.1638	.3581	423	99.673
8	.3632	.6691	775	126.68
9	.3142	.5578	695	123.24
A	.2718	.4285	776	163.63
B	.4347	.7976	931	154.56
C	.3429	.8021	591	112.05
D	.3869	.6347	714	150.37
E	.444	.7976	867	172.45
F	.363	.6457	706	152.39
G	.3454	.8021	718	149.23
H	.4119	.7518	862	164.01
I	.3479	.8593	491	121.92
J	.2968	.4312	540	123.59
K	.3703	.7077	843	145.87
L	.3141	.5114	573	141.09
M	.2726	.4516	960	213.13
N	.3462	.6532	912	169.35
O	.3384	.8672	678	122.34
P	.3436	.6457	694	142.67
Q	.2974	.8672	881	151.22
R	.3778	.7556	861	143.36
S	.4012	.5863	703	134.94
T	.3466	.8246	632	156.01
U	.3683	.6532	671	149.45
V	.2389	.3737	608	135.79
W	.2986	.5545	960	204.12
X	.332	.6082	754	126.63
Y	.3492	.6082	591	117.93
Z	.4068	.616	672	144.86

Table B4
Font-dependent visual properties: Times new roman.

Character	Average visual overlap: Pixel	Competitor distance: Pixel	Number of pixels	Perimetric complexity
0	.3122	.6085	1102	81.488
1	.216	.8768	601	57.977
2	.3195	.5576	856	79.784
3	.3017	.5866	788	73.503
4	.2535	.4541	1017	71.859

(continued on next page)

Table B4 *(continued)*

Character	Average visual overlap: Pixel	Competitor distance: Pixel	Number of pixels	Perimetric complexity
5	.2984	.5866	835	71.301
6	.2922	.6068	1014	83.321
7	.2694	.5576	654	71.119
8	.3488	.6664	1070	89.427
9	.2958	.6085	1018	82.423
A	.2587	.4541	1084	103.53
B	.4157	.7051	1642	108.28
C	.3089	.745	1014	95.386
D	.3613	.7025	1578	97.876
E	.4019	.7918	1220	146.43
F	.3561	.7937	1016	114
G	.3049	.7514	1327	113.06
H	.3525	.689	1560	140.6
I	.2565	.8768	706	63.66
J	.2712	.624	778	55.966
K	.3491	.666	1485	115.6
L	.3428	.7918	932	85.528
M	.2032	.3706	1860	172.03
N	.3103	.6119	1292	139.36
O	.2942	.9156	1380	87.253
P	.3523	.7937	1156	86.745
Q	.277	.9156	1617	97.798
R	.3893	.666	1497	103.35
S	.3818	.6664	1148	92.386
T	.288	.785	922	98.483
U	.3147	.683	1137	125
V	.2611	.5877	1008	90.281
W	.2811	.3928	1742	147.75
X	.3267	.6119	1277	112.68
Y	.3104	.5877	1017	90.274
Z	.3523	.5395	1155	104.65

Table B5
Font-invariant visual properties.

	Average visual overlap: Feature	Competitor distance: Feature	Number of features
0	.096	1	1
1	.1423	.3519	3
2	.1144	.2963	3
3	.1008	.5556	2
4	.1481	.3519	3
5	.136	.2778	3
6	.0766	.5	2
7	.1367	.4444	2
8	.0865	.5	2
9	.0778	.5	2
A	.1091	.2778	3
B	.1409	.5556	3
C	.0698	.4167	1
D	.1534	.4167	2
E	.1744	.5139	4
F	.1772	.4815	3
G	.1149	.3333	3
H	.18	.4722	3
I	.1747	.5278	3
J	.1429	.3333	3
K	.1149	.2407	3
L	.1802	.5	2
M	.073	.3333	4
N	.1479	.4074	3
O	.096	1	1
P	.1558	.4444	2
Q	.0861	.5833	2
R	.131	.3333	3
S	.0853	.4167	2
T	.1746	.5278	2
U	.145	.4074	3
V	.075	.4167	2
W	.0714	.3333	4
X	.0821	.4167	2

(continued on next page)

Table B5 *(continued)*

	Average visual overlap: Feature	Competitor distance: Feature	Number of features
Y	.0837	.2778	3
Z	.1507	.4444	3

References

Abboud, S., Maidenbaum, S., Dehaene, S., Amedi, A., 2015. A number-form area in the blind. Nat. Commun. 6, 1–9. http://dx.doi.org/10.1038/ncomms7026.
Allison, T., McCarthy, G., Nobre, A., Puce, A., Belger, A., 1994. Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. Cereb. Cortex 4 (5), 544–554 , (Retrieved from)⟨http://www.ncbi.nlm.nih.gov/pubmed/7833655⟩.
Arguin, M., Fiset, S., Bub, D., 2002. Sequential and parallel letter processing in letter-by-letter dyslexia. Cogn. Neuropsychol. 19 (6), 535–555. http://dx.doi.org/10.1080/02643290244000040.
Beeli, G., Esslen, M., Jäncke, L., 2007. Frequency correlates in grapheme-color synaesthesia. Psychol. Sci. 18 (9), 788–792. http://dx.doi.org/10.1111/j.1467-9280.2007.01980.x.
Benford, F., 1938. The law of anomalous numbers. Proc. Am. Philos. Soc. 78, 551–572.
Benoit, L., Lehalle, H., Molina, M., Tijus, C., Jouen, F., 2013. Young children's mapping between arrays, number words, and digits. Cognition 129 (1), 95–101. http://dx.doi.org/10.1016/j.cognition.2013.06.005.
Bernard, M.L., Chaparro, B.S., Mills, M.M., Halcomb, C.G., 2003. Comparing the effects of text size and format on the readibility of computer-displayed Times New Roman and arial text. Int. J. Hum. Comput. Stud. 59 (6), 823–835. http://dx.doi.org/10.1016/S1071-5819(03)00121-6.
Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., Zorzi, M., 2010. Numerical estimation in preschoolers. Dev. Psychol. 46 (2), 545–551. http://dx.doi.org/10.1037/a0017887.
Boles, D.B., Clifford, J.E., 1989. An upper- and lowercase alphabetic similarity matrix, with derived generation similarity values. Behav. Res. Methods, Instrum., Comput. 21 (6), 579–586. http://dx.doi.org/10.3758/BF03210580.
Briggs, R., Hocevar, D., 1975. A new distinctive feature theory for upper case letters. J. Gen. Psychol. 93, 87–93.
Brunsdon, R., Coltheart, M., Nickels, L., 2006. Severe developmental letter-processing impairment: a treatment case study. Cogn. Neuropsychol. 23 (6), 795–821. http://dx.doi.org/10.1080/02643290500310863.
Bub, D.N., Arguin, M., Lecours, A.R., 1993. Jules Dejerine and his interpretation of pure alexia. Brain Lang. 45, 531–559.
Butterworth, B., 2005. The development of arithmetical abilities. J. Child Psychol. Psychiatry Allied Discip. 46 (1), 3–18. http://dx.doi.org/10.1111/j.1469-7610.2004.00374.x.
Cantlon, J.F., Pinel, P., Dehaene, S., Pelphrey, K.A., 2011. Cortical representations of symbols, objects, and faces are pruned back during early childhood. Cereb. Cortex 21 (1), 191–199. http://dx.doi.org/10.1093/cercor/bhq078.
Caramazza, A., Hillis, A.E., 1990. Levels of representation, co-ordinate frames, and unilateral neglect. Cogn. Neuropsychol. 7 (5/6), 37–41 , (Retrieved from)⟨http://coglab.wjh.harvard.edu/~caram/PDFs/1990_Caramazza_Hillis.pdf⟩.
Castles, A., Coltheart, M., 2004. Is there a causal link from phonological awareness to success in learning to read? Cognition 91 (1), 77–111. http://dx.doi.org/10.1016/S0010-0277(03)00164-1.
Castles, A., Coltheart, M., Wilson, K., Valpied, J., Wedgwood, J., 2009. The genesis of reading ability: what helps children learn letter-sound correspondences? J. Exp. Child Psychol. 104 (1), 68–88. http://dx.doi.org/10.1016/j.jecp.2008.12.003.
Chanceaux, M., Grainger, J., 2012. Serial position effects in the identification of letters, digits, symbols, and shapes in peripheral vision. Acta Psychol. 141 (2), 149–158. http://dx.doi.org/10.1016/j.actpsy.2012.08.001.
Chang, Y.-N., Furber, S., Welbourne, S., 2012. Modelling normal and impaired letter recognition: implications for understanding pure alexic reading. Neuropsychologia 50 (12), 2773–2788. http://dx.doi.org/10.1016/j.neuropsychologia.2012.07.031.
Changizi, M.A., Shimojo, S., 2005. Character complexity and redundancy in writing systems over human history. Proc. Biol. Sci./R. Soc. 272 (1560), 267–275. http://dx.doi.org/10.1098/rspb.2004.2942.
Chochon, F., Inserm, U., 1999. Differential contributions of the left and right inferior parietal lobules to number processing. J. Cogn. Neurosci. 11 (6), 617–630.
Cipolotti, L., 1995. Multiple routes for reading words, why not numbers? Evidence from a case of arabic numeral dyslexia. Cogn. Neuropsychol. 12 (3), 313–342. http://dx.doi.org/10.1080/02643299508252001.
Cohen, D.J., 2009. Integers do not automatically activate their quantity representation. Psychon. Bull. Rev. 16 (2), 332–336. http://dx.doi.org/10.3758/PBR.16.2.332.
Cohen, L., Dehaene, S., 1995. Number processing in pure alexia: the effect of hemispheric asymmetries and task demands. Neurocase 1 (2), 121–137. http://dx.doi.org/10.1080/13554799508402356.
Collis, N.L., Kohnen, S., Kinoshita, S., 2013. The role of visual spatial attention in adult developmental dyslexia. Q. J. Exp. Psychol. 66 (2), 245–260. http://dx.doi.org/10.1080/17470218.2012.705305.
Dehaene, S., 1992. Varieties of numerical abilities. Cognition 44 (1–2), 1–42. http://dx.doi.org/10.1016/0010-0277(92)90049-N.
Deloche, G., Seron, X., 1987. Numerical Transcoding: a general production model. In: Mathematical Disabilities: A Cognitive Neuropsychological Perspective 1. Lawrence Erlbaum Associates, Hillsdale, NJ, 137–170.
Duñabeitia, J.A., Dimitropoulou, M., Grainger, J., Hernández, J.A., Carreiras, M., 2012. Differential sensitivity of letters, numbers, and symbols to character transpositions. J. Cogn. Neurosci. 24 (7), 1610–1624. http://dx.doi.org/10.1162/jocn_a_00180.
Egeth, H., Jonides, J., Wall, S., 1972. Parallel processing of multielement displays. Cogn. Psychol. 3 (4), 674–698. http://dx.doi.org/10.1016/0010-0285(72)90026-6.
Finkbeiner, M., Coltheart, M., 2009. Letter recognition: from perception to representation. Cogn. Neuropsychol. 26 (1), 1–6. http://dx.doi.org/10.1080/02643290902905294.
Fiset, D., Blais, C., Ethier-Majcher, C., Arguin, M., Bub, D., Gosselin, F., 2008. Features for identification of uppercase and lowercase letters. Psychol. Sci. 19 (11), 1161–1168. http://dx.doi.org/10.1111/j.1467-9280.2008.02218.x.
Ganor-Stern, D., Tzelgov, J., 2008. Across-notation automatic numerical processing. J. Exp. Psychol. Learn., Mem., Cogn. 34 (2), 430–437. http://dx.doi.org/10.1037/0278-7393.34.2.430.
García-Orza, J., Perea, M., Muñoz, S., 2010. Are transposition effects specific to letters? Q. J. Exp. Psychol. 63 (8), 1603–1618. http://dx.doi.org/10.1080/17470210903474278.
García-Orza, J., Perea, M., Abu Mallouh, R., Carreiras, M., 2012. Physical similarity (and not quantity representation) drives perceptual comparison of numbers: evidence from two Indian notations. Psychon. Bull. Rev. 19 (2), 294–300. http://dx.doi.org/10.3758/s13423-011-0212-8.
Grainger, J., Hannagan, T., 2014. What is special about orthographic processing? Writ. Lang. Lit. 17 (2), 225–252. http://dx.doi.org/10.1075/wll.17.2.03gra.
Grainger, J., Bertrand, D., Lété, B., Beyersmann, E., Ziegler, J.C., 2016a. A developmental investigation of the first-letter advantage. J. Exp. Child Psychol. 152, 161–172. http://dx.doi.org/10.1016/j.jecp.2016.07.016.
Greenblatt, S.H., 1973. Alexia without agraphia or hemianopsia. Brain 96, 307–316.
Grossi, D., Fragassi, N. a., Orsini, a., De Falco, F.A., Sepe, O., 1984. Residual reading capability in a patient with alexia without agraphia. Brain Lang. 23 (2), 337–348 , (Retrieved from)⟨http://www.ncbi.nlm.nih.gov/pubmed/6518359⟩.
Grotheer, M., Ambrus, G.G., Kovács, G., 2016a. Causal evidence of the involvement of the number form area in the visual detection of numbers and letters. NeuroImage 132, 314–319. http://dx.doi.org/10.1016/j.neuroimage.2016.02.069.
Grotheer, M., Herrmann, K.-H., Kovacs, G., 2016b. Neuroimaging evidence of a bilateral representation for visually presented numbers. J. Neurosci. 36 (1), 88–97. http://dx.doi.org/10.1523/JNEUROSCI.2129-15.2016.
Hamilton, J.P., Mirkin, M., Polk, T. a., 2006. Category-level contributions to the alphanumeric category effect in visual search. Psychon. Bull. Rev. 13 (6), (Retrieved from)⟨http://www.ncbi.nlm.nih.gov/pubmed/17484438⟩.
Hammond, E.J., Green, D.W., 1982. Detecting targets in letter and non-letter arrays. Can. J. Psychol. 36 (1), 67–82.
Hickok, G., 2013. The functional neuroanatomy of language. Handb. Clin. Neurophysiol. 10 (3), 61–70. http://dx.doi.org/10.1016/B978-0-7020-5310-8.00003-X.
Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92 (1–2), 67–99. http://dx.doi.org/10.1016/j.cognition.2003.10.011.
Hofstadter, D., McGraw, G., 1995. Letter Spirit: esthetic perception and creative play in the rich microcosm of the Roman alphabetFluid Concepts and Creative Analogies. Basic Books, New York, NY, 407–466.
Holender, D., Peereman, R., 1987. Differential processing of phonographic and logographic single-digit numbers by the two hemispheres. In: Deloche, G., Seron, X. (Eds.), Mathematical Disabilities: A Cognitive Neuropsychological Perspective. Lawrence Erlbaum Associates, Hillsdale, NJ, 43–86.
Hurst, M., Anderson, U., Cordes, S., 2016. Mapping among number words, numerals, and non-symbolic quantities in preschoolers. J. Cogn. Dev. 8372 (September). http://dx.doi.org/10.1080/15248372.2016.1228653.
Ingles, J.L.J.L., Eskes, G.A., 2008. A comparison of letter and digit processing in letter-by-letter reading. J. Int. Neuropsychol. Soc. 14 (1), 164–173 , (Retrieved from) ⟨http://journals.cambridge.org/abstract_S1355617708080119⟩.
James, K.H., James, T.W., Jobard, G., Wong, A.C.-N., Gauthier, I., 2005. Letter processing in the visual system: different activation patterns for single letters and strings. Cogn., Affect., Behav. Neurosci. 5 (4), 452–466 , (Retrieved from)⟨http://www.ncbi.nlm.nih.gov/pubmed/16541814⟩.
Jones, M.N., Mewhort, D.J.K., 2004. Case-sensitive letter and bigram frequency counts from large-scale English corpora. Behav. Res. Methods, Instrum., Comput. 36 (3), 388–396. http://dx.doi.org/10.3758/BF03195586.
Jonides, J., Gleitman, H., 1972. A conceptual category effect in visual search: o as letter or as digit. Percept. Psychophys. 12 (6), 457–460. http://dx.doi.org/10.3758/BF03210934.

Katz, R.B., Sevush, S., 1989. Positional Dyslexia. Brain Lang. 37, 266–289.
Kinoshita, S., Lagoutaris, S., 2010. Priming by NUMB3R5 does not involve top-down feedback. J. Exp. Psychol. Learn., Mem., Cogn. 36 (6), 1422–1440. http://dx.doi.org/10.1037/a0020609.
Kinoshita, S., Robidoux, S., Mills, L., Norris, D., 2013. Visual similarity effects on masked priming. Mem. Cogn.. http://dx.doi.org/10.3758/s13421-013-0388-4.
Kinoshita, S., Robidoux, S., Guilbert, D., Norris, D., 2015. Context-dependent similarity effects in letter recognition. Psychon. Bull. Rev., 1–7. http://dx.doi.org/10.3758/s13423-015-0826-3.
Landerl, K., Fussenegger, B., Moll, K., Willburger, E., 2009. Dyslexia and dyscalculia: two learning disorders with different cognitive profiles. J. Exp. Child Psychol. 103 (3), 309–324. http://dx.doi.org/10.1016/j.jecp.2009.03.006.
Lanthier, S.N., Risko, E.F., Stolz, J.A., Besner, D., 2009. Not all visual features are created equal: early processing in letter and word recognition. Psychon. Bull. Rev. 16 (1), 67–73. http://dx.doi.org/10.3758/PBR.16.1.67.
Larsen, J., Baynes, K., Swick, D., 2004. Right hemisphere reading mechanisms in a global alexic patient. Neuropsychologia 42 (11), 1459–1476. http://dx.doi.org/10.1016/j.neuropsychologia.2004.04.001.
Marinus, E., Mostard, M., Segers, E., Schubert, T.M.T.M., Madelaine, A., Wheldall, K., 2016. A special font for people with dyslexia: does it work and, if so, why? Dyslexia 22 (3), 233–244. http://dx.doi.org/10.1002/dys.1527.
Martin, A., Schurz, M., Kronbichler, M., Richlan, F., 2015. Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. Hum. Brain Mapp. 36 (5), 1963–1981. http://dx.doi.org/10.1002/hbm.22749.
Mason, M., 1982. Recognition time for letters and nonletters: effects of serial position, array size, and processing order. J. Exp. Psychol.: Hum. Percept. Perform. 8 (5), 724–738 , (Retrieved from)⟨http://www.ncbi.nlm.nih.gov/pubmed/6218232⟩.
Mayzner, M.S., Tresselt, M.E., 1965. Tables of single-letter and digram frequency counts for various word-length and letter-position combinations. Pyschonomic Monogr. Suppl. 3 (2), 13–32.
McClelland, J.L., Rumelhart, D.D.E., 1981. An interactive activation model of context effects in letter perception: part 1. An account of basic findings. Psychol. Rev. 88, 375–407 , (Retrieved from)⟨http://psycnet.apa.org/journals/rev/88/5/375/⟩.
McCloskey, M., 1992. Cognitive mechanisms in numerical processing: evidence from acquired dyscalculia. Cognition 44 (1–2), 107–157. http://dx.doi.org/10.1016/0010-0277(92)90052-J.
McCloskey, M., Schubert, T.M., 2014. Shared versus separate processes for letter and digit identification. Cogn. Neuropsychol. 0 (0), 1–24. http://dx.doi.org/10.1080/02643294.2013.869202.
Merkley, R., Wilkey, E.D., Matejko, A.A., 2016. Exploring the origins and development of the visual number form area: a functionally specialized and domain-specific region for the processing of number symbols? J. Neurosci. 36 (17), 4659–4661. http://dx.doi.org/10.1523/JNEUROSCI.0710-16.2016.
Miozzo, M., Caramazza, A., 1998. Varieties of pure alexia: the case of failure to access graphemic representations. Cogn. Neuropsychol. 15 (1), 203–238. http://dx.doi.org/10.1080/026432998381267.
Moret-Tatay, C., Perea, M., 2011. Do serifs provide an advantage in the recognition of written words? J. Cogn. Psychol. 23 (5), 619–624. http://dx.doi.org/10.1080/20445911.2011.546781.
Mueller, S.T., Weidemann, C.T., 2012. Alphabetic letter identification: effects of perceivability, similarity, and bias. Acta Psychol. 139, 19–37. http://dx.doi.org/10.1016/j.actpsy.2011.09.014.
New, B., Grainger, J., 2011. On letter frequency effects. Acta Psychol. 138 (2), 322–328. http://dx.doi.org/10.1016/j.actpsy.2011.07.001.
Norris, D., Kinoshita, S., 2012. Reading through a noisy channel: why there's nothing special about the perception of orthography. Psychol. Rev. 119 (3), 517–545. http://dx.doi.org/10.1037/a0028450.
Norris, D., Kinoshita, S., van Casteren, M., 2010. A stimulus sampling theory of letter identity and order. J. Mem. Lang. 62 (3), 254–271. http://dx.doi.org/10.1016/j.jml.2009.11.002.
Park, J., Hebrank, A., Polk, T.A., Park, D.C., 2011. Neural dissociation of number from letter recognition and its relationship to parietal numerical processing. J. Cogn. Neurosci. 24 (1), 39–50. http://dx.doi.org/10.1162/jocn_a_00085.
Patterson, K., Wilson, B., 1990. A rose is a rose or a nose: a deficit in initial letter identification. Cogn. Neuropsychol. 7 (5/6), 447–477.
Pelli, D.G., Burns, C.W., Farell, B., Moore-Page, D.C., 2006. Feature detection and letter identification. Vis. Res. 46 (28), 4646–4674. http://dx.doi.org/10.1016/j.visres.2006.04.023.
Perea, M., Dun, J.A., Carreiras, M., 2008. R34D1NG W0RD5 W1TH NUMB3R5. J. Exp. Psychol.: Hum. Percept. Perform. 34 (1), 237–241. http://dx.doi.org/10.1037/0096-1523.34.1.237.
Perri, R., Bartolomeo, P., Silveri, M.C., 1996. Letter dyslexia in a letter-by-letter reader. Brain Lang. 53 (3), 390–407. http://dx.doi.org/10.1006/brln.1996.0055.
Peters, L., De Smedt, B., Op de Beeck, H.P., 2015. The neural representation of Arabic digits in visual cortex. Front. Hum. Neurosci. 9 (October), 517. http://dx.doi.org/10.3389/fnhum.2015.00517.
Petit, J.-P., Grainger, J., 2002. Masked partial priming of letter perception. Vis. Cogn. 9 (3), 337–353. http://dx.doi.org/10.1080/13506280042000207.
Piazza, M., Eger, E., 2015. Neural foundations and functional specificity of number representations. Neuropsychologia 83, 257–273. http://dx.doi.org/10.1016/j.neuropsychologia.2015.09.025.
Pitchford, N.J., Ledgeway, T., Masterson, J., 2008. Effect of orthographic processes on letter position encoding. J. Res. Read. 31 (1), 97–116. http://dx.doi.org/10.1111/j.1467-9817.2007.00363.x.
Polk, T.A., Farah, M.J., 1994. Late experience alters vision. Nature 376, 648–649 , (Retrieved from)⟨http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=40359&tool=pmcentrez&rendertype=abstract⟩.
Polk, T.A., Farah, M.J., 1995. Brain localization for arbitrary stimulus categories: a simple account based on Hebbian learning. Proc. Natl. Acad. Sci. 92 (26), 12370–12373 , (Retrieved from)⟨http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=40359&tool=pmcentrez&rendertype=abstract⟩.
Polk, T.A., Farah, M.J., 1998. The neural development and organization of letter recognition: evidence from functional neuroimaging, computational modeling, and behavioral studies. Proc. Natl. Acad. Sci. 95, 847–852.
Polk, T.A., Stallcup, M., Aguirre, G.K., Alsop, D.C., Esposito, M.D., Detre, J.A., Farah, M.J., 2002. Neural specialization for letter recognition. J. Cogn. Neurosci. 14 (2), 145–159. http://dx.doi.org/10.1162/089892902317236803.
Prado, J., Mutreja, R., Zhang, H., Mehta, R., Desroches, A.S., Minas, J.E., Booth, J.R., 2011. Distinct representations of subtraction and multiplication in the neural systems for numerosity and language. Hum. Brain Mapp. 32 (11), 1932–1947. http://dx.doi.org/10.1002/hbm.21159.
Price, G.R., Ansari, D., 2011. Symbol processing in the left angular gyrus: evidence from passive perception of digits. NeuroImage 57 (3), 1205–1211. http://dx.doi.org/10.1016/j.neuroimage.2011.05.035.
Rath, D., Domahs, F., Dressel, K., Claros-Salinas, D., Klein, E., Willmes, K., Krinzinger, H., 2015. Patterns of linguistic and numerical performance in aphasia. Behav. Brain Funct. 11 (2). http://dx.doi.org/10.1186/s12993-014-0049-1.
Reicher, G.M., 1969. Perceptual recognition as a function of meaninfulness of stimulus material. J. Exp. Psychol.: Gen. 81 (2), 275–280. http://dx.doi.org/10.1037/h0027768.
Reinke, K., Fernandes, M., Schwindt, G., O'Craven, K., Grady, C.L., 2008. Functional specificity of the visual word form area: general activation for words and symbols but specific network activation for words. Brain Lang. 104, 180–189. http://dx.doi.org/10.1016/j.bandl.2007.04.006.
Rosa, E., Perea, M., Enneson, P., 2016. The role of letter features in visual-word recognition: evidence from a delayed segment technique. Acta Psychol. 169, 133–142. http://dx.doi.org/10.1016/j.actpsy.2016.05.016.
Rothlein, D., Rapp, B.C., 2014. The similarity structure of distributed neural responses reveals the multiple representations of letters. NeuroImage 89, 331–344. http://dx.doi.org/10.1016/j.neuroimage.2013.11.054.
Roux, F.-E., Lubrano, V., Lauwers-Cances, V., Giussani, C., Démonet, J.-F., 2008. Cortical areas involved in Arabic number reading. Neurology 70 (3), 210–217. http://dx.doi.org/10.1212/01.wnl.0000297194.14452.a0.
Sanocki, T., Dyson, M.C., 2012. Letter processing and font information during reading: beyond distinctiveness, where vision meets design. Atten., Percept. Psychophys. 74 (1), 132–145. http://dx.doi.org/10.3758/s13414-011-0220-9.
Schubert, T.M., Badcock, N.A., Kohnen, S., 2016. Children's processing of identity and position in letter, digit, and symbol strings. Manuscript submitted for publication.
Schubert, T.M., McCloskey, M., 2013. Prelexical representations and processes in reading: evidence from acquired dyslexia. Cogn. Neuropsychol. 30 (6), 360–395. http://dx.doi.org/10.1080/02643294.2014.880677.
Shalev, R.S., Gross-Tsur, V., 1993. Developmental dyscalculia and medical assessment. J. Learn. Disabil. 26 (2), 134–137. http://dx.doi.org/10.1016/B978-0-12-801871-2.00007-1.
Shum, J., Hermes, D., Foster, B.L., Dastjerdi, M., Rangarajan, V., Winawer, J., Parvizi, J., 2013. A brain area for visual numerals. J. Neurosci. 33 (16), 6709–6715. http://dx.doi.org/10.1523/JNEUROSCI.4558-12.2013.
Simpson, I.C., Mousikou, P., Montoya, J.M., Defior, S., 2012. A letter visual-similarity matrix for Latin-based alphabets. Behav. Res. Methods. http://dx.doi.org/10.3758/s13428-012-0271-4.
Sinclair, A., Siegrist, F., Sinclair, H., 1983. Young children's ideas about the written number system. In: Rogers, D.R., Sloboda, J.A. (Eds.), The Acquisition of Symbolic Skills. Plenum Press, New York, NY, 535–542.
Smilek, D., Carriere, J.S. a., Dixon, M.J., Merikle, P.M., 2007. Grapheme frequency and color luminance in grapheme-color synaesthesia. Psychol. Sci. 18 (9), 793–795. http://dx.doi.org/10.1111/j.1467-9280.2007.01981.x.
Solso, R.L., King, J.F., 1976. Frequency and versatility of letters in the English language. Behav. Res. Methods Instrum. 8 (3), 283–286.
Starrfelt, R., Behrmann, M., 2011. Number reading in pure alexia-A review. Neuropsychologia 49, 2283–2298. http://dx.doi.org/10.1016/j.neuropsychologia.2011.04.028.
Starrfelt, R., Habekost, T., Leff, A.P., 2009. Too little, too late: reduced visual span and speed characterize pure alexia. Cereb. Cortex 19 (12), 2880–2890. http://dx.doi.org/10.1093/cercor/bhp059.
Starrfelt, R., Habekost, T., Gerlach, C., 2010. Visual processing in pure alexia: a case study. Cortex 46 (2), 242–255. http://dx.doi.org/10.1016/j.cortex.2009.03.013.
Starrfelt, R., Lindegaard, M., Bundesen, C., 2015. Confusing confusability: on the problems of using psychophysical measures of letter confusability in neuropsychological research. Cogn. Neuropsychol. July, 1–7. http://dx.doi.org/10.1080/02643294.2015.1061488.
Taylor, D.A., 1978. Identification and categorization of letters and digits. J. Exp. Psychol.: Hum. Percept. Perform. 4 (3), 423–439. http://dx.doi.org/10.1037/0096-1523.4.3.423.
Taylor, J.S.H., Rastle, K., Davis, M.H., 2013. Distinct neural specializations for learning to read words and name objects. J. Cogn. Neurosci. 26 (9), 2128–2154. http://dx.doi.org/10.1162/jocn.
Teng, E.L., Sperry, R.W., 1973. Interhemispheric interaction during simultaneous bilateral presentation of letters or digits in commissurotomized patients. Neuropsychologia 11 (2), 131–140. http://dx.doi.org/10.1016/0028-3932(73)90001-8.
Thesen, T., McDonald, C.R., Carlson, C., Doyle, W., Cash, S., Sherfey, J., Halgren, E.,

2012. Sequential then interactive processing of letters and words in the left fusiform gyrus. Nat. Commun. 3, 1284. http://dx.doi.org/10.1038/ncomms2220.

Tydgat, I., Grainger, J., 2009. Serial position effects in the identification of letters, digits, and symbols. J. Exp. Psychol.: Hum. Percept. Perform. 35 (2), 480–498. http://dx.doi.org/10.1037/a0013027.

Walker, P., Hinkley, L., 2003. Visual memory for shape-colour conjunctions utilizes structural descriptions of letter shape. Vis. Cogn. 10 (8), 987–1000. http://dx.doi.org/10.1080/13506280344000185.

Wheeler, D.D., 1970. Processes in word recognition. Cogn. Psychol. 1 (1), 59–85. http://dx.doi.org/10.1016/0010-0285(70)90005-8.

Wong, B., Szücs, D., 2013. Single-digit Arabic numbers do not automatically activate magnitude representations in adults or in children: evidence from the symbolic same-different task. Acta Psychol. 144 (3), 488–498. http://dx.doi.org/10.1016/j.actpsy.2013.08.006.

Ziegler, J.C., Pech-Georgel, C., Dufau, S., Grainger, J., 2010. Rapid processing of letters, digits and symbols: what purely visual-attentional deficit in developmental dyslexia? Dev. Sci. 13 (4). http://dx.doi.org/10.1111/j.1467-7687.2010.00983.x.