The role of allograph representations in font-invariant letter identification

David Rothlein[a,*] and **Brenda Rapp**[b,c,d]

[a]Department of Veterans Affairs, VA Boston Healthcare System, Boston Massachusetts

[b]Department of Cognitive Science, Johns Hopkins University, Baltimore MD 21218, USA

[c]Department of Neuroscience, Johns Hopkins University, Baltimore MD 21218, USA

[d]Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore MD 21218, USA

Abstract

The literate brain must contend with countless font variants for any given letter. How does the visual system handle such variability? One proposed solution posits stored structural descriptions of basic letter shapes that are abstract enough to deal with the many possible font variations of each letter. These font-invariant representations, referred to as allographs in this paper, while frequently posited, have seldom been empirically evaluated. The research reported here helps to address this gap with two experiments that examine the possible influence of allograph representations on visual letter processing. In these experiments, participants respond to pairs of letters presented in an atypical font in two tasks—visual similarity judgments (Experiment 1) and same/different decisions (Experiment 2). By using Representational Similarity Analysis in conjunction with Linear Mixed Effect Models (RSA-LMEM) we show that the similarity structure of the responses to the atypical font is influenced by the predicted similarity structure of allograph representations even after accounting for font-specific visual shape similarity. Similarity due to symbolic (abstract) identity, name, and motor representations of letters are also taken into account providing compelling evidence for the unique influence of allograph representations in these tasks. These results provide support for the role of allograph representations in achieving font-invariant letter identification.

TRANSLATIONAL ABSTRACT

In everyday reading, we often encounter unfamiliar letter shapes such as when we read handwritten notes or read in a font we have not seen before. Although our brains may have never before processed these shapes we are, nonetheless, usually able to effortlessly recognize them. The research reported here is directed at furthering our understanding of how this happens. In two experiments, research participants made decisions regarding the visual characteristics of letters presented in a highly unusual font. We found evidence that one important part of the letter recognition process involves automatically generating mental information about the standard,

---

[*]Corresponding author. Department of Veterans Affairs, VA Boston Healthcare System. 150 S. Huntington Ave. Boston, MA 02130. david.rothlein@gmail.com (D. Rothlein).

canonical geometric shapes of letters. These results have implications for reading teachers and clinicians working with acquired or developmental reading impairments as understanding the types of mental representations that skilled readers use for efficient and effective letter recognition can be used to structure teaching and remediation to optimize learning.

Keywords

Letter representation; Allographs; Symbolic Letter Identities (SLIs); Abstract Letter Identities (ALIs); Representational Similarity Analysis (RSA); Linear Mixed Effects Modeling (LMEM)

Introduction

Like many other visual objects, letters are often easily identified despite large differences in size, position, and shape. While certain large differences in the visual shapes of letters can be irrelevant (ear or ear or EAR), small visual differences can be highly significant (lend me your ear vs. lend me your car). To identify relevant visual specificity in the face of immense stimulus variability, many theories of reading assume that letter recognition is accomplished through the computation of increasingly abstract representations. In letter recognition, one type of representation that is often posited involves the spatial-geometric representation of each letter's canonical (font-invariant) shape(s). These representational types are sometimes referred to as "allograph" representations, corresponding to what, in the object recognition literature, have been referred to as structural or canonical "descriptions" (Marr & Nishihara, 1978). Given that many, if not most, theories of letter recognition and reading posit a key role for letter representations of this type (Brunsdon, Coltheart, & Nickels, 2006; Chauncey, Holcomb, & Grainger, 2008; Cox, Coueignoux, Blesser, & Eden, 1982; Dehaene, Cohen, Sigman, & Vinckier, 2005; Grainger, Rey, & Dufau, 2008; Herrick, 1975; Hofstadter & McGraw, 1995; Walker & Hinkley, 2003; Wong & Gauthier, 2007), it is surprising that, outside of cognitive neuropsychological literature, there has been relatively little research examining their existence. In this paper we report on two behavioral experiments carried out with neurotypical participants that provide evidence that allograph representations play a role in letter recognition.

Levels of representation in letter identification

Figure 1 schematizes key representational types that are assumed in many theories of letter recognition and reading. Of these, we will focus specifically on: computed letter shapes, allograph representations and abstract (symbolic) letter identities. (Note that this focus does not imply that there are not additional important representations involved in letter recognition).

At the earliest stages of visual processing, internal representations are largely isomorphic with the pattern of light energy transduced by the visual sensory processes. These representations encode image properties such as intensity values at specific retinotopic positions that can be approximated by differences in pixel configurations. According to many prominent views of visual-spatial processing (see Riesenhuber & Poggio, 1999 for review), these low-level retinotopic representations undergo a series of transformations in

which the stimulus is represented in terms of increasingly more complex visual features. For example, at an early stage, the stimulus may be represented as a set of oriented bars, that would serve as input to simple visual feature detectors (e.g., right angles, crosses and curves). Subsequently, the lines, curves, and angles that compose the low-level visual features are integrated into a unitary shape at the level of **computed stimulus-shape representations**. The term "computed" highlights that these are not stored shape representations but rather that they are computed for any stimulus shapes, regardless of whether the viewer has previously encountered them. In the context of letter processing, computed stimulus-shape representations are font specific. In other words, there are different computed stimulus-shape representations for **a** and **a**.

Further along the visual processing pathway and at a higher level of visuospatial abstraction, **allographs** are font-invariant, spatial/geometric structural descriptions of canonical letter-forms that have been learned and stored in long-term memory (Schubert & McCloskey, 2013). Thus, the allograph representation of **a** would be active in response to a range of computed stimulus-shape representations that share its defining spatial/geometric features, including computed shape representations for font variants such as **a** and **a**. These various visual shapes are considered font-specific tokens or exemplars of the same allograph. Thus, a key feature of allographs is that they are font-invariant—equally activated by different font-specific stimulus-shape representations. In contrast, **a**, a, **A** and **N** correspond to four different allographs, each with different spatial/geometric features. The stored information about the font-invariant spatial features represented by allographs allows viewers to determine the shape category that a given letter stimulus corresponds to. It is typically assumed that normalization processes and font-specific translation rules (Gauthier, Wong, Hayward, & Cheung, 2006; Hofstadter & McGraw, 1995; Sanocki & Dyson, 2012; Walker, 2008) are applied to computed stimulus shape representations to allow for the correct categorization of stimulus shapes into their corresponding allographs, resulting in font-invariant recognition of letter shapes.

In addition to font-invariant recognition of familiar letter shapes, we are also able to recognize that **A**, **a** and **a**—despite corresponding to different allographs—all correspond to the same letter identity [A], while **N** corresponds to a different letter identity. Font, shape and case invariant representations of letter identity are generally proposed in most theories of letter recognition and reading (e.g., Besner, Coltheart, & Davelaar, 1984; Brunsdon et al., 2006; Petit, Midgley, Holcomb, & Grainger, 2006) and are commonly referred to as abstract letter identities (ALI). However, given the many ambiguities surrounding the term "abstract", and to emphasize that these representations are not visuo-spatial but rather entirely abstract (i.e., lacking in modality-specific content), we will refer to them as **symbolic letter identities (SLIs)**. As indicated in Figure 1, SLIs provide the format used in word reading and allow for the transcoding between the different modality-specific representations of letters, such as allographs, motor plans and phonological letter names. While SLIs do not form the focus of the research reported in this paper, the research will also provide evidence regarding the role of SLIs in letter identification.

Evidence of allograph representations

To date, evidence for allograph representations has come primarily from neuropsychological case studies (Brunsdon et al., 2006; Caplan & Hedley-Whyte, 1974; Chanoine, Ferreira, Demonet, Nespoulous, & Poncet, 1998; Dalmás & Dansilio, 2000; Miozzo & Caramazza, 1998; Rapp & Caramazza, 1989; Schubert & McCloskey, 2013). For example, case GV (Miozzo & Caramazza, 1998), following a stroke, performed perfectly on a letter decision task which involved deciding for each letter and pseudoletter stimulus whether the shape corresponded to a real letter or not. This type of task requires accessing stored representations that differentiate real from pseudoletters and, within the model depicted in Figure 1, this decision could be made based on either allograph or SLI representations. To identify which of these representations were available to GV, Miozzo and Caramazza (1998) had her perform tasks that required access to SLIs such as cross-case matching (e.g., indicating whether letter pairs such as **a** and **A** correspond to the same letter). Miozzo and Caramazza (1998) found that, despite her flawless performance in letter decision tasks, GV was highly impaired in cross-case matching (accuracy 63%). On this basis, the authors ruled out the that GV's good performance in the letter decision task could be based on SLIs and concluded that, instead, it relied on intact allograph representations (for a similar set of findings, see also, Rapp & Caramazza, 1989). In other work, Schubert & McCloskey (2013) specifically examined whether or not allograph representations are font invariant. They reported on case LHD whose performance on letter decision and cross-case matching, was generally similar to GV's, and, therefore, indicated disruption in access to SLI's. They also tested her on a cross-*font* matching task (e.g., do **a** and **a** correspond to the same letter?) which, according to the account depicted in Figure 1, should be able to be successfully carried out by someone who only has access to allograph representations. As predicted, LHD was perfect on this task (accuracy 100%) and they concluded that LHD's pattern of performance can be understood by positing allograph representations that are invariant to font differences. These allograph representations supported her accurate letter decision and cross-font matching in the face of impaired access to SLIs.

There have been few behavioral studies with neurotypical individuals that have investigated the distinction between font-invariant allograph representations and SLIs. In fact, in the only study we are aware of addressing this issue, Walker and Hinkley (2003) reported evidence for a level of letter representation that is consistent with allographs in a study examining the memory for color-letter associations. Walker and Hinkley (2003) found that, when articulatory suppression techniques were employed to disrupt participants from phonologically recoding letter and color names, participants were better at remembering color-letter associations when the letter itself was colored on a white background vs. when the letter was white on a colored background. Critical to the issue of allograph representation, they found that this benefit for color-letter association generalized at test for letters that differed in font but that it failed to generalize across differences in case. This effect was observed even when certain changes in font were judged to be more visually dissimilar than changes in case, demonstrating that the effect was not solely based on the degree of visual similarity between letter forms. These results were interpreted as demonstrating that the colors were linked to letter representations that were abstract enough to survive changes in font but not abstract enough to survive changes in case—in other

words, the colors were linked to allograph representations (which Walker & Hinkley referred to as structural descriptions). The results are compelling, however due to the strong visual memory component, it is not clear to what degree paired-associations tap normal letter recognition processes. Therefore, exploring the issue using a different approach could provide convergent evidence in support of allograph representations.

Neuroimaging studies investigating the existence of allograph representations are limited as well. Studies have found cross-font priming effects (e.g., Gauthier et al., 2000; Qiao et al., 2010) for single letters or words presented in different fonts in the left fusiform gyrus, an area associated with letter and word processing in reading (Cohen et al., 2000) and spelling (Rapp & Lipka, 2011). However, the challenge in interpreting these results involves identifying the specific type of representation driving the cross-font priming effects. Since single letter stimuli were used in the Gauthier et al. (2000) study, priming of SLI representations or even low-level visual representations could have provided the basis for the observed cross-font priming effects. Word stimuli were used in the Qiao et al. (2010) study and, therefore, while the priming effects they report could have originated in allograph representations, they also could have originated in SLIs, lexical, semantic, or even phonological representations that are shared by words that differ only in font.

Alternative views: Are allograph representations necessary?

Letter and word recognition models that do not include allograph representations have also been proposed. For example, exemplar-based models (Tenpenny, 1995) rely on large storage capacities that encode memory traces of every letter exemplar that a person has viewed. When a letter stimulus is viewed, it is identified by computing the similarity of the stimulus to each of the stored exemplars. The identity of the stimulus is determined by assuming that it corresponds to the label (e.g., the name) of the nearest stored exemplar (see Goldinger, 1998 for an exemplar model of spoken word recognition). Note that these models fall within the general category of grounded and embodied cognitive theories that rely solely on sensory and motor representations (Barsalou, 2008; Tulving, 1983; Wilson, 2002). Empirical evidence consistent with exemplar, episodic or instance-based accounts of visual word recognition comes largely from priming effects that are specific to the surface features such as font (Sanocki, 1992) or episodic context (Carlson, Alejano, & Carr, 1991; Grant & Logan, 1993) and on long-term priming for pseudowords (Grant & Logan, 1993). Tenpenny (1995) presents a review of these findings and arguments (but see Bowers (2000) and Marsolek (2004) for critiques of purely episodic models of reading).

Also of relevance is the fact that there are computational models that learn to carry out handwritten letter recognition by learning to associate a range of letter shapes with the same letter (e.g., Graves et al., 2009). The classification performance of these models can be impressive and is achieved without explicit coding of allograph representations. However, these networks are sufficiently opaque and complex that the nature of the internal representations that actually develop during learning are not understood, leaving open the possibility that, in fact, allograph-like representations may develop and play a significant computational role. Furthermore, there is no guarantee that the computations that underlie these models' success are the same that enable letter recognition in humans.

Representational similarity analysis (RSA) and the challenges of isolating allograph representations

The goal of this study is to determine if allograph representations are automatically deployed in letter identification. Representational similarity analysis is one approach to identifying the representational types that contribute to some dependent measure. While the approach has been widely applied to neuroimaging studies (Kriegeskorte, Mur, & Bandettini, 2008), the same principles have also been applied to behavioral data (e.g., Shepard & Chipman, 1970; Watson, Akins, & Enns, 2012). The basic assumption of the approach is that if a particular representational type (e.g., allograph) is involved in a task (e.g., letter identification) then processing items that are similar with regard to that representational type should produce similar behaviors (or neural responses). For example, letter pairs that correspond to similar allographs should be expected to be rated as similar in a similarity judgment task, take longer to discriminate in a same/different task, etc. In other words, RSA allows one to determine the involvement of one or more representational types in task performance by testing whether the similarity structure predicted by a representational type is present in some dependent measure (see Figure 2). In this way, RSA is based on the evaluation of second order isomorphisms, asking: Is the similarity structure of the data space correlated with the similarity structure of the representational space of interest? In this particular context, we would want to determine if the structure of the allograph similarity space is reflected in behavioral responses to letter stimuli. With regard to allograph representations, one could develop a similarity matrix for a set of letter pairs based on their allographic similarity and then statistically evaluate the extent to which this predicted representational similarity matrix (pRSM) explains the variance of behavioral responses to those letters pairs in the observed representational similarity matrix (oRSM). This approach allows us to evaluate the hypothesis of interest: If allograph representations are involved in letter recognition then we expect a significant and unique contribution of the allograph pRSM in explaining the variance in the oRSM. On the other hand, if allograph representations are not involved in letter recognition, then the allograph pRSM will not explain significant variance in the oRSM.

However, there are important challenges in determining whether or not allograph representations are involved in letter identification. The key challenge is to distinguish between the predicted contributions of computed stimulus-shape representations and allograph representations (see Figure 1). This is challenging because the similarity space for computed stimulus-shape representations would usually be highly similar to the similarity space for allographs. For example, if viewers judge **d** and **D** to be more visually similar than **f** and **H**, is this because the computed stimulus shapes for **d** and **D** are more similar than the stimulus shapes for **f** and **H** or, alternatively, because their allograph representations are more similar? In other words, if most letter pairs that are considered to be similar in terms computed stimulus-shape representations also have similar allograph representations, then it will be difficult to distinguish one representational type from the other. As we have noted above, the failure to distinguish between these two representational types has been a shortcoming of some of the previous work on this topic and, for this reason, the current study is specifically designed to address this challenge.

Second, there are the challenges of developing computed stimulus-shape and allograph pRSMs. Key to the RSA analysis approach is developing distinct representational similarity matrices that represent stimulus-shape similarity and allograph similarity. These are both difficult given the lack of clear understanding of these representational levels. For example, for the allograph pRSM, although we assume that allographs are font-invariant, spatial/geometric descriptions of letter-forms, their specific representational content and format has not been determined. (Note that this is also the case for the analogous structural descriptions posited to play a similar role in object recognition, where this topic is an active area of research (e.g. Davitt, Cristino, Wong, & Leek, 2014)).

Meeting the challenges: The current study

The key innovation of this study that allows us to address the challenge of distinguishing between the similarity spaces of computed stimulus-shape and allograph representations is the use of a stimulus font that is highly atypical. The fact that these atypical letter stimuli (although recognizable as letters) are markedly different in shape from letters in a more standard font, allows for the possibility that computed letter-shape and allograph representations may be distinguishable. Following Hofstadter and McGraw (1995) we call the novel font a gridfont and examples of letters in the font that are used in this study are shown in Figure 3. As can be seen, some letters in the gridfont that are highly similar would not be highly similar in a standard font and, by assumption, would not be highly similar in terms of their allograph representations. This should allow for dissociations between the similarity structure at the level of computed letter-shapes and the level of allographs that may be sufficient to isolate the influence of each these representational types via an RSA approach.

With regard to the challenge of developing a suitable stimulus-shape similarity matrix (pRSM) for the atypical gridfont letters, while there are many approaches to estimating pairwise stimulus-shape similarity, all are limited by the fact that we lack adequate theories of the content of these representations. Common approaches to estimate visual similarity in the absence of such a theory include computing pixelwise similarity, empirically derived similarity measures based on visual confusions (for review see Mueller & Weidemann, 2012), or visual similarity judgments (Boles & Clifford, 1989; Rothlein & Rapp, 2014; Simpson, Mousikou, Montoya, & Defior, 2012). We sought an approach, such as visual similarity judgments, that captures visual similarity above the level of the pixel-like representation but which would not be influenced by letter recognition and higher level letter representations. To do so, we extended an approach used by Lupyan et al. (2010). As a control for the visual similarity between B-b, these researchers used the pair B-p reasoning that b and p differ only by a horizontal reflection. On this basis (see also, Egeth & Blecker, 1971), we developed a computed letter-shape pRSM for all of the gridfont letter pairs based on visual similarity judgments obtained from participants presented with pairs of *rotated* gridfont letters (see Figure 3) which they were not able to recognize as letters. In this way, the rotated gridfont letters provided a match to the upright gridfont letters in terms of computed stimulus shapes as they are visually identical except for the rotation which prevents viewers from identifying the stimuli as letter and, therefore, prevents the influence

of high level representations associated with letter recognition[1]. This approach yields a pRSM that encodes estimated pairwise computed stimulus-shape similarity.

Finally, with regard to developing an allograph similarity matrix, given our limited understanding of the representational content of structural descriptions, we made the assumption that allograph representations resemble typical (canonical) font letter shapes (such as A and B). Thus, we assumed that the visual similarity of letters viewed in a typical font constitutes a good proxy for the similarity structure of allograph representations (see Figure 3). We consider this to be a reasonable assumption because: a) allograph representations are learned from experience and therefore, because typical fonts are those that we have experienced the most, typical fonts should have the greatest influence on the content of allograph representations and b) it is computationally sensible that allograph tuning to typical and frequent letter forms would maximize efficiency during letter recognition at many visual levels, including font-to-allograph translation (Walker, 2008). On this basis, the similarity space of allographs (the allograph pRSM) can be derived empirically from any of a number of tasks that allow one to compute similarity values for pairs of letters viewed in a typical font. Specifically, in this study the allograph pRSM was developed from the pairwise similarity judgments produced by participants viewing pairs of letters in a fairly standard font (Consolas).

In sum, the use of an atypical gridfont allows for the possibility of dissociable stimulus-shape and allograph representations. Based on the assumptions and innovations described just above, the experimental approach we adopted was as follows. We used two tasks that involved the visual presentation of letter pairs in an atypical gridfont—visual similarity judgments (Experiment 1) and same/different judgments (Experiment 2). These tasks provided the dependent measures (ratings, RTs and accuracies) used for constructing three *observed* representational similarity matrices (oRSMs), one for each dependent measure. We then used the Linear Mixed Effects Modeling (LMEM) regression approach (Bates, Mächler, Bolker, & Walker, 2014) to determine the unique contribution of various representational types to the similarity structure of the oRSMs. To do so, we used as regressors the *predicted* representational similarity matrices (pRSMs) for the following representational types: computed stimulus-shape similarity, allograph similarity, symbolic letter identity, motor stroke similarity, and phonetic similarity of letter names. In the methods section, we describe in detail how these different pRSMs were developed. The application an RSA approach to data obtained from viewing letters in a novel font allowed us to address the following question: When individuals view and recognize letters in an

---

[1]One possible objection to the use of rotated stimuli as a visual control for the of upright letters is that computed shape similarity may interact with object orientation. For example, it is possible that the computed shape similarity for **p q** is different from that of the structurally-matched pair **b d** (e.g., shared inward facing top-sided loop features may appear more visually similar than shared inward-facing bottom-sided loop features). However, if a shape-by-orientation interaction effect exists it is likely to be small. This is supported by findings from Konkle et al., (2010) in which distinctiveness judgments for sets of object images were shown to predict localization times. Importantly, shape distinctiveness judgments were elicited using both upright and inverted sets of the same objects. While the shape-by-orientation interaction account would predict that the upright distinctiveness judgments would be better predictors of search times since the objects in both tasks shared the same orientation, the distinctiveness judgments to the inverted objects actually had a slight numerical advantage in predicting the search times for the upright objects. Whether or not interactions between orientation and visual similarity exist is an interesting avenue for further research that could have consequences for a number of studies that have used rotated letters or objects as visual controls to distinguish the influence of perceptual and conceptual features (e.g., letters: Egeth & Blecker, 1971 and Lupyan et al., 2010; objects: Konkle et al., 2010 or scenes: Naber & Nakayama, 2013).

atypical font, does the letter recognition process recruit allograph representations that were learned based on past experiences with more typical fonts?

Experiment 1 – Pairwise visual similarity judgments of letters

Visual similarity judgments to pairs of letters were collected for three stimulus sets from three groups of participants in order to address the following question: Do allograph representations explain any unique variance in visual similarity judgments to letters presented in an atypical gridfont even after taking into account stimulus-shape similarity and several other possible dimensions of letter similarity?

The first group was presented letters in an atypical upright gridfont. The responses from this group served as the dependent measure (the oRSM) in the regression (LMEM) analysis developed to test for the role of allograph (and other) representations in visual similarity judgments. The results from groups 2 and 3 were used to generate pRSMs for the regression analysis of the data from the first group. The second group was presented with rotated versions of letters in the atypical gridfont. Because these letters were difficult to recognize as letters, the responses to these letters provided an estimate of the visual letter-shape similarity structure (pRSM) of the atypical upright gridfont, estimating computed letter-shape similarity. Finally, a third group was presented with letters in a typical font. Responses to these letters served as an estimate for the similarity structure (pRSM) of allograph representations.

Methods

Each experiment was run on Amazon's Mechanical Turk (AMT) and coded using HTML and JavaScript. JavaScript code made use of jquery1.8.3 (https://ajax.googleapis.com/ajax/libs/jquery/1.8.3/jquery.min.js) and TimTurkTools (Tim Brady - https://timbrady.org/turk/TimTurkTools.js) packages.

Participants—153 participants were recruited from Amazon's Mechanical Turk (AMT). Participants were instructed not to participate in any experiment if they had a history of reading or spelling disabilities. Furthermore, they were instructed not to participate if they were literate in any other written script besides the Roman alphabet although there was no way to independently verify this. Participants were only recruited from the US and were paid $1.00 for their participation. The 153 participants were divided into 3 groups based on the stimuli used: the Upright Gridfont, the Rotated Gridfont, and the Typical Font. For the Upright Gridfont Group, 54 participants were recruited. Participants had to have participated in at least 100 HITs (tasks on AMT) with an approval rating of at least 95%. For the Rotated Gridfont Group, 50 participants were recruited. Participants had to have participated in at least 1000 HITs with an approval rating of at least 90%. For the Typical Font Group, 49 participants were recruited. Participants had to have participated in at least 1000 HITs with an approval rating of at least 90%.

Stimuli and procedures

<u>Group 1: Upright Gridfont:</u> Figure 4 (a) depicts the letter-stimuli (7 different letter identities) presented to the Upright Gridfont Group. In order to control for visual cues like size and curvature, all stimuli were constructed from a limited feature-set of straight lines and were matched to be the same size (see Figure 4 (d) for more details). The limited feature set was used to ensure that features like stimulus orientation (e.g. r, *r*) and line thickness would not serve as cues indicating different shapes. Furthermore, since stimuli were matched in height and width, stimulus size could not serve as a cue to letter case or identity. Finally, the Upright Gridfont was designed to have low-level visual characteristics that substantially differ from more typical and frequent fonts.

Each experimental stimulus was centered within a 100px by 100px stimulus space which was vertically centered within a 500px (width) by 600px (height) frame with a black border. During the stimulus familiarization portion of the experiment, one stimulus was presented centered within the display frame. The written feedback displaying the correct response was centered as well. During the similarity judgment portion of the experiment, 2 letter stimuli were displayed side by side (centered within their own stimulus space). The center of each stimulus space was 150px away from a fixation dot that was centered both vertically and horizontally within the display frame.

Participants first completed a familiarization task for which they were instructed to indicate the identity of each letter stimulus by pressing the appropriate key on their keyboard. They were also told responses were not case-sensitive. A stimulus character would appear within the display frame. Once a response was given (correct or incorrect), the correct answer appeared on the screen in Arial font (e.g., "`lower-case g`"). Participants began the next trial by pressing spacebar. Each stimulus appeared twice. The stimulus familiarization portion consisted of 66 trials, displaying each of the Upright Gridfont stimuli twice.

In the experimental task, participants were instructed to rate the visual similarity of each stimulus pair on a scale of 1 through 5 by pressing 1, 2, 3, 4, or 5 on their keyboards, with 1 indicating low similarity and increasing numbers indicating increasing similarity up to 5. A reminder of this scale remained visible throughout the task as well as a countdown of the number of remaining trials. Once the task began, a stimulus pair was shown. A keyboard response immediately triggered the appearance of the next stimulus pair. Similarity judgments and reaction times were recorded. The similarity judgment portion consisted of 22 practice trials and 528 experimental trials. The 528 experimental trials consisted of 1 trial for each possible different-stimulus pairs for the set of 33 stimuli—including five digit stimuli not analyzed further in this paper. The 22 practice trials were randomly selected from the set of possible different-stimulus pairs. Each letter in a pair was randomly assigned a left or right position on each trial and trial order was randomized for each worker.

<u>Group 2: Rotated Gridfont:</u> Experimental stimuli and design were identical to those used with Group 1 except for three differences. First and most importantly, the stimuli consisted of the gridfont presented to Group 1 except each stimulus was flipped around its vertical axis and rotated 90°ccw[2] (Figure 4, b). Next, there was no familiarization portion of the experiment and the stimuli were referred to as shapes instead of letters. Finally, at the end of

the experiment, a survey was presented to each participant asking if any of the shapes were recognized and if so to provide a few examples. Like the Upright Gridfont Group, the similarity judgment portion consisted of 22 practice trials and 528 experimental trials. The 528 experimental trials consisted of every possible different-stimulus pair from the set of 33 rotated stimuli. The 22 practice trials were randomly selected from the set of possible different-stimulus pairs. Stimulus position (left or right) was randomized for each trial and trial order was randomized for each participant.

Group 3: Typical Font: The experimental design was identical to that used with Group 2 except for the stimulus set (Figure 4 c). The characters were presented upright in the font Consolas. The similarity judgment portion consisted of 22 practice trials and 255 experimental trials. The 255 experimental trials consisted of every possible different-stimulus pair within the set of 23 stimuli. The 22 practice trials were randomly selected from the set of possible different-stimulus pairs. Stimulus position (left or right) was randomized for each trial and trial order was randomized for each worker.

Data analysis

Removing outliers: Given that data were collected via Mechanical Turk, it was especially important to identify outlier participants. This is because it became apparent that some participants, being unmonitored, randomly pressed buttons to complete the task as quickly as possible. To identify and remove such participants in a principled way, the following procedure was employed. First, for each participant, a similarity-to-sample value was computed by correlating (Pearson) a given participant's set of similarity judgments to each letter pair with each of the other participants' set of similarity judgments, resulting in 53 r values. The similarity-to-sample value consisted of the mean of these r values. This was computed for each participant resulting in 54 similarity-to-sample values. Finally, outliers were defined as participants whose similarity-to-sample values fell 1.5 standard deviations below the mean similarity-to-sample value (0.3032). For Group 1, eight participants were found to be outliers in this manner and removed from further analysis. The mean similarity to sample was increased to 0.3926 after it was recomputed with the remaining 46 participants. For Group 2, seven outlier participants were identified and removed (n = 43) following the same procedure described above, increasing the mean similarity to sample value from 0.32 to 0.42 and for Group 3, four outlier participants were identified and removed (n = 45) increasing the mean similarity to sample value from 0.23 to 0.27. Finally, for Group 1, trials where the response times were faster than 200ms or slower than a minute were removed. A total of 1.1% of trials were removed in this step.

Measuring representational influence: A Linear Mixed Effects Modeling approach: We combined Representational Similarity Analysis (RSA) and linear mixed effects modeling (LMEM) ("lme4" library, version 1.1–12 (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2015)) to quantify the influence of allograph representations on each participant's visual similarity judgments of the *Upright Gridfont* letters while accounting for a number of visual and non-visual controls. The dependent variable consisted of the

---

[2]For simplicity, these stimuli will be referred to as rotated characters.

similarity judgments for each pair of *Upright Gridfont* letters from each participant in Group 1. Seven fixed-effect predictors were included in this model, five of which consisted of pRSMs which are described below. The two additional fixed effect predictors were response RTs for each similarity judgment (log-transformed to address positive skew) and trial order. In terms of random effects, we included random intercepts by item and by participant and 5 random slopes by participant—one for each of the pRSM-based fixed effects. Slopes were not computed for trial order or response RTs as we were not interested in drawing conclusions from the fit of these variables.

The Five pRSMs: (1) **Rotated Gridfont pRSM**: the results from Group 2 were compiled into a single group pRSM. This was done by normalizing each participant's responses to a mean of 0 and a standard deviation of 1. Then, the participants' responses were averaged together so each rotated pair had a single mean similarity value. This group pRSM provided an empirical estimate of the stimulus shape similarity of the *Upright Gridfont*. (2) **Typical Font pRSM**: the results from Group 3 were combined into a single group pRSM using the same procedure as the *Rotated Gridfont pRSM*. This pRSM approximated similarity at the level of allograph representations. (3) **Letter Identity pRSM:** all letter pairs with the same identity (e.g., **a**, **A**) were assigned 1 and all others a 0. (4) **Phonetic Feature pRSM:** using the set of phonetic features corresponding to each letter name (features taken from an interactive IPA phonetic feature chart http://www.linguistics.ucsb.edu/projects/featuresoftware/index.php), phonetic feature overlap was computed for each letter pair by summing the features from each letter name that overlap with the features from the other letter name and dividing that sum by the total number of features across the pair of letter names. (5) **Motoric pRSM:** this was adapted from a bi-stroke feature similarity metric (Wiley et. al., 2015) that was based on a feature set validated against written letter confusions produced by individuals with acquired dysgraphia (Rapp & Caramazza, 1997).

Results

Upright gridfont letter familiarization task (Group 1)—While overall accuracy on the familiarization task administered to Group 1was quite high (mean = 92%), four stimuli had accuracies below 70% and were removed from subsequent analysis. For the remaining 25 stimuli (shown in Figure 4 a), the distribution of accuracies revealed that they were highly recognizable with accuracies greater than 90%.

Survey results for Group 2—Participants in Group 2 were asked if the rotated gridfont stimuli looked familiar and if so, to give examples of what they were. Of the 44 participants, only 7 identified multiple rotated stimuli as alphabetic letters. These 7 participants were removed from all subsequent analyses[3]. We did not remove 3 participants who either reported a single letter (e.g. "t shape") or only reported letters that were not actually included (e.g., "the letter M"). A majority of responses were statements indicating that the rotated stimuli were not familiar (23 participants) and the remaining 11 participants provided

---

[3]The analyses were additionally performed including these 7 participants and the results were substantively the same (i.e., no insignificant results became significant or vice versa).

non-alphanumeric examples (e.g., pencil, house, fish). These results indicated that rotating the gridfont rendered the stimuli unrecognizable for the vast majority of participants.

Similarity matrices—Figure 5 depicts the group averages of the pairwise similarity values obtained from the three groups, corresponding to the group Upright Gridfront oRSM, the group Rotated Gridfont pRSM (estimating computed stimulus-shape similarity) and the group Typical Font pRSM (estimating allograph similarity). The group RSMs were formed by normalizing each participant's responses to a mean of 0 and a standard deviation of 1 and then averaging the normalized responses to each letter-pair across participants. The figure highlights that there was a similar, but not identical pattern of responses to the Upright and Rotated Gridfonts indicating that while much of the variance in the similarity judgments to the Upright Gridfont letters can be explained by computed stimulus-shape similarity, at least some of the unexplained variance might be accounted for by assuming the influence stored letter-representations (e.g., allograph representations) on the visual similarity judgments. In addition, Table 1 reports the pairwise correlations between each RSM. As expected the Upright Gridfont correlates most highly with the Rotated Gridfont (r = 0.78). This indicates that the Rotated Gridfont provides a good measure of visual similarity for the Upright Gridfont. Importantly, the Rotated and Upright Gridfonts differ in their correlations with the Typical Font with the correlation of the Upright Gridfont (r = 0.69) being greater than the Rotated Gridfont (r = 0.40), suggesting the possible additional influence of allograph representations in the processing of Upright Gridfont letters. A similar pattern is found when comparing the Upright and Rotated Gridfont correlations to the Letter Identity pRSM (r = 0.47 and r = 0.20 respectively). Finally, no such differences between the Upright and Rotated Gridfont RSMs were observed when correlating them to the phonetic or motor pRSMs.

Linear Mixed Effects RSA—Using a measure of R^2 proposed by Nakagawa and Schielzeth (2013), we can report the conditional R^2 (R^2_c) that provides a measure of the variance explained by both the fixed and random effects combined. Additionally, the R^2_m value provides a measure of the variance explained by the fixed effects. The R^2_c and R^2_m values were computed in R using the 'MuMIn' package version 1.15.6 (Barton, 2016). The $R^2_c = 0.58$, suggesting that a considerable amount of the trial-to-trial variance in the visual similarity judgments for the pairs of Upright Gridfont letters can be explained by the effects included in the model and the $R^2_m = 0.22$ suggested that nearly a quarter of the variance can be explained by the fixed-effects. We interpret the beta coefficients (β) and *t* values associated with the beta coefficients as indicating the degree of influence of each of the fixed effects. To evaluate the statistical significance of the *t* values, the number of degrees of freedom along with the p value associated with each beta weight *t* were estimated in R using the 'lmerTest' package version 2.0–32 (Kuznetsova et al., 2016). As indicated in Table 2, four of the five pRSMs were found to make significant contributions to the visual similarity judgement values obtained from Group 1: Upright Gridfont. Indexing the role of allograph processing, the Typical Font pRSM was found to exert a unique and positive influence on similarity judgments (β = 0.21, p<0.001). Also exerting an influence on the pattern of response times were the Rotated Gridfont pRSM (β = 0.34, p<0.001), the Letter Identity

pRSM (β = 0.08, p<0.05) and the Phonetic Feature pRSM (β = −0.09, p<0.001). Only the Motoric pRSM did not exert a significant influence with (β = −0.03, p=0.17).

Discussion of experiment 1

The results of Experiment 1 demonstrate that similarity judgments to the atypical gridfont letters are influenced by the visual similarity of the letter-pairs presented in a more typical font. In other words, letter pairs that are relatively visually similar in a typical font will be judged as more similar when viewed in the gridfont. Critically, this effect cannot be reduced to the influences of computed stimulus-shape similarity (estimated from the similarity judgments from the Group 2 judgments of the Rotated Gridfont), symbolic letter identity, phonetic letter name similarity, or motoric similarity. We interpret the influence of the typical font similarity on judgments involving Upright Gridfont letters as resulting from the activation of stored allograph representations when the Upright Gridfont letters are processed. It is also noteworthy that the Letter Identity pRSM also positively and uniquely influenced the Upright Gridfont similarity judgments indicating that symbolic letter identities are also accessed and influence visual similarity judgments in this task.

The negative direction of the influence of the Phonetic Feature pRSM on the Upright Gridfont similarity judgments indicates that letter-pairs that have a greater degree of phonetic feature overlap are judged to be less visually similar than letters that have fewer shared phonetic features. The interpretation is not obvious. One possibility is that it could be due the over-application of an explicit strategy to reduce the influence non-visual letter representations from the visual similarity judgments by judging letters that are similar on a non-visual dimension (e.g. letter name) as being less similar than one would judge them otherwise. However, the fact that this negative affect is observed for phonetic letter name similarity while other non-visual representations like symbolic letter identity exert a positive influence suggests this is may not be an inadequate explanation.

Experiment 2 – Physical same-different decision task

That stored allographs influenced visual similarity judgments (above and beyond the effects of computed stimulus shape similarity) was quite apparent from the results in Experiment 1, although the precise nature of this influence was not entirely clear. For example, instead of an unconscious influence on the visual judgments, it could be that participants misinterpreted the instructions that the similarity judgments were to be based strictly on visual criterion and, instead, judged letter similarity more generally, considering other dimensions of similarity. In that case, it would be less surprising that stored letter knowledge would make a contribution to decision times. Furthermore, since visual similarity was never explained to the participants, different participants could have used different criteria when mentally computing visual similarity. Finally, visual similarity judgments occur on a relatively slow timescale that may allow the influence of stored letter representations to influence responses. Therefore, a task where decisions are quicker and less explicitly based on similarity may be a better measure of automatic letter processing. We address these issues in Experiment 2, collecting reaction times and accuracies for same-different judgments to gridfont letter pairs. Additionally, since the primary finding in Experiment 1—that allograph

representations significantly influenced visual similarity judgments—is a novel finding, convergence across multiple tasks would serve to strengthen confidence in this conclusion.

In this experiment, two letter images were shown and each participant simply decided whether the two stimulus images were visually identical or not. Two LMEMs were developed to predict RT (model 1) and the accuracy (model 2) for the responses to pairs of different letters. Since the same-different decisions were to be based on visual identity/non-identity decisions, a participant who misunderstood the directions would be easy to spot since they would systematically respond incorrectly when the letter pair consisted of font or case variants of the same letter identity.

Methods

Participants—100 participants were recruited from Amazon's Mechanical Turk (AMT). Participants were instructed not to participate if they had a history of reading or spelling disabilities or if they were literate in any other written script besides the Roman alphabet. Neither of these conditions was verified. Participants were only recruited from the US and had to have participated in at least 1000 HITs (tasks on AMT) with an approval rating of at least 95%. They were split into two groups based on the type of stimuli shown. 50 participants were shown the upright gridfont and were paid $1.00 for their participation. The other 50 participants were shown the rotated gridfont and paid $0.80 for their participation

Stimuli and procedure—Experimental stimuli consisted of a subset[4] of the Gridfont stimuli used in Experiment 1, presented in the upright orientation to one group (*Group 1: Upright Gridfont*) and the rotated orientation to the other group (*Group 2: Rotated Gridfont*). For the *Upright Gridfont Group*, the stimulus familiarization portion consisted of 38 trials, displaying each stimulus twice. For both groups, the visual same-different decision portion consisted of 15 practice trials and 285 experimental trials. The 285 experimental trials consisted of every possible different-stimulus pair within the set of 19 stimuli (171 trials or 60% of total) and 6 repetitions of each of the 19 possible same pairs (114 trials or 40% of total). The 15 practice trials were randomly selected from the set of possible experimental trials. Stimulus position (left or right) was randomized for each trial and trial order was randomized for each participant. The familiarization task was carried out only with Group 1 and in the same way as described in Experiment 1.

The experimental task collected visual same or different judgments for simultaneously presented stimuli for both Groups 1 and 2 (Figure 6). Participants were informed through written instruction that they would see two shapes and a dot on the center of the screen and that they were to decide whether the two shapes are visually identical or not. Pressing **s** on the keyboard indicated a "same" response and **d** indicated a "different" response. They were instructed to respond with their first impression as quickly as possible. A countdown of the number of remaining trials remained visible throughout the task. Once a trial began, a fixation dot appeared in isolation for either 400ms or 800ms (143 and 142 trials respectively, randomly assigned) followed by a pair of stimuli. The keyboard response triggered the

---

[4]Six letter stimuli, corresponding to the [D] or [H] letter identities, were removed from the stimulus set for Experiment 2 in order to keep the experimental duration shorter than 20 minutes.

appearance of a hyperlink which participants clicked to begin the next trial, allowing for self-pacing. Accuracy and reaction times (from stimulus onset to keyboard response) were recorded. For the Rotated Gridfont Group there was no familiarization portion of the experiment and the experiment was followed by the same survey presented to Experiment 1, Group 2.

Data analysis

Removing outliers—Only RTs and errors from trials in which different letters were presented are analyzed below. For each participant, a different trial's RT was considered to be an outlier and removed if it fell outside 2.5 standard deviations (above or below) that participant's mean different trial RT. Response times from incorrect responses were also removed.

Outlier participants were removed using the same similarity-to-sample procedure described in Experiment 1. The 50 similarity to sample values were considerably lower than in the similarity judgment experiments suggesting greater cross-participant variability in this task. Outliers were defined as participants whose similarity to sample value fell more than 1.5 standard deviations below the mean similarity to sample value. In Group 1: Upright Gridfont, 4 participants were identified as outliers and removed from further analysis, increasing the mean similarity-to-sample from 0.066 to 0.08. In Group 2: Rotated Gridfont, 3 participants were identified as outliers and removed, increasing the mean similarity-to-sample from 0.065 to 0.073.

LMEM-RSA—Following a similar method as in Experiment 1, LMEM-RSA was used to examine the trial-to-trial performance of the different responses and test for the influence of allograph representations. Unlike Experiment 1, two models (RT and accuracy) were examined.

In Model 1, RTs (log-normalized) from Group 1: Upright Gridfont served as the dependent variable with 7 fixed effect predictors—5 pRSMs, trial order and the RT from the previous trial[5] (on the first trial, the average RT was used) (Baayen & Milin, 2010). The 5 pRSMs were the same as Experiment 1 with one important exception, the **Rotated Gridfont pRSM** was composed of the RTs to the different responses from Experiment 2's Group 2 (instead of the Experiment 1's Group 2 that had carried out the visual similarity judgment task). Specifically, each participant's RTs to correct different responses were normalized to a mean of 0 and a standard deviation of 1. Then, the RTs for each different letter pair were averaged across participants to form a group Rotated Gridfont pRSM. Note that the Typical Font pRSM was the same as the one used in Experiment 1 (based on Experiment 1's Group 3 performance) as we assumed that those data would serve to index allograph similarity regardless of task. We included 3 random effects: random intercepts by items and by participants and random slopes by participants for the Typical Font pRSM. Slopes of the other pRSMs were not modeled because the model failed to converge when they were included.

---

[5]RTs have a high degree of temporal autocorrelation that is not stimulus item driven. Including RTs from the previous trial controls for this source of RT variation (Baayen & Milin, 2010).

Model 2 used response accuracy as the dependent variable and the analysis was performed using generalized linear mixed models for binomial data. The six fixed effect predictors in this model were trial order and the 5 pRSMs. For the **Rotated Gridfont pRSM** the total number of errors produced by Group 2 participants for each different letter pair was used (instead of the average z-scored RT used in Experiment 1). With regard to random effects, because errors were so rare, only intercepts were modeled by participants and item.

Results

Letter familiarization task—Overall accuracy on the familiarization task was very high (mean = 98%). All of the letter stimuli used in this experiment were highly recognizable with accuracies greater than 90% (Range: 91%–100%).

Similarity matrices—Figure 7 depicts the group RSMs for RTs to different trials for the Upright (a) and Rotated (b) Gridfont groups. The group RSMs were generated by normalizing each participant's RTs to a mean of 0 and a standard deviation of 1 and then averaging the normalized responses to each letter-pair across participants. The figure highlights that there is a similar, but not identical pattern of responses to the Upright and Rotated Gridfonts suggesting that while some of the variance in the RTs to the Upright Gridfont letters can be explained by computed stimulus-shape similarity (indexed by the Rotated Gridfont RSM), there is unexplained variance. The focus of this experiment is in determining whether a significant amount of this variance can be accounted for by assuming the influence of allograph representations, whose similarity structure is approximated by similarity judgments to letters pairs presented in the Typical Font (see Figure 5 c). Figure 8 makes a similar point depicting error totals (trials on which participants gave "same" responses to" trials with different letter pairs) for the Upright (a) and Rotated Gridfont (b) groups. In addition, Table 3 reports the pairwise correlations between each pRSM and the group RT RSMs. Both the Rotated Gridfont and the Typical Font pRSMs correlate most strongly with the Upright Gridfont (r = 0.56 for both). This indicates that both the Rotated Gridfont and the Typical Font provide a good measure of visual similarity for the Upright Gridfont. Importantly, the Typical Font correlates less with the Rotated Gridfont (r = 0.42) than with the Upright Gridfont (r=0.56) suggesting that allograph representations influence visual same-different decision RTs for letter pairs presented in Upright Gridfont. A similar pattern can be found when comparing the Upright and Rotated Gridfont correlations to the Letter Identity pRSM (r = 0.46 and r = 0.21 respectively). Finally, no such differences between the Upright and Rotated Gridfont RSMs were observed when correlating them to the phonetic or motor pRSMs.

Table 4 reports the same pair-wise correlations for the error counts instead of the RTs. The most notable difference with the RT RSMs is the weaker correlation between the Upright and Rotated Gridfont RSMs (r = 0.12 compared to 0.56) suggesting that, compared to RT, accuracy is less influenced by low-level stimulus-shape similarity than it is by higher level letter identity representations (e.g., allographs and symbolic letter identity). There were substantial differences in the correlations between both the Typical Font and Letter Identity RSMs and the Upright and Rotated Gridfonts revealing an influence of allograph and SLI representations in error likelihood (Typical Font pRSM: Upright Gridfont r = 0.47, Rotated

Gridfont r = 0.16; Letter Identity pRSM: Upright Gridfont r = 0.42, Rotated Gridfont r = 0.18).

Survey results—Participants were asked if the rotated character stimuli looked familiar and if so, to give examples. Of the 47 participants, only 13 identified them as including multiple alphabetic letters. These participants were removed from further analyses. Three additional participants recognized the letter T and failed to correctly identify anything else and were not removed. As for Experiment 1, the majority of responses were statements indicating that the rotated stimuli were not familiar (20 participants) and the remaining 11 participants provided non-alphanumeric examples (e.g. cross, angles, house). These results demonstrate again that rotating the gridfont rendered the stimuli unrecognizable to a majority of participants.

LMEM-RSA – Model 1: Response times—For Model 1, an $R^2_c = 0.49$ indicated that about half of the trial-to-trial variance of the RTs for the different pairs of Upright Gridfont letters could be explained by the variables included in the model. A considerably smaller proportion of the variance than in Experiment 1 could be explained by the fixed effects ($R^2_m = 0.05$). As in Experiment 1, we interpreted the beta coefficients (β) and t values associated with the beta coefficients as indicating the degree of influence of each of the fixed effects on the response to make same/different judgments for different pairs (see Table 5 for details). Indexing the influence of allograph representations, the Typical Font pRSM exerted a unique and positive influence on RTs with $\beta = 0.011$ and p<0.05, indicating that gridfont letter pairs that are visually similar in a typical font have slower RTs for than letter pairs that are not. With regard to computed stimulus-shape similarity, there was a significant and positive effect of Rotated Gridfont RT pRSM with $\beta = 0.025$, p<0.001), as well as of symbolic letter identity with Letter Identity pRSM with $\beta = 0.012$, p<0.01. Neither the Phonetic Feature pRSM ($\beta = -0.004$, p=0.37) nor the Motoric pRSM ($\beta = -0.004$, p=0.29) exerted a significant influence on same/different response times.

LMEM-RSA – Accuracy—For Model 2, $R^2_c = 0.05$ indicating a relatively small[6] proportion of the error variance could be explained by this model. This is perhaps due to the fact that the model was trying to predict a relatively few number of errors (232 errors total, comprising 3% of the different trials included in the model). Despite this, Typical Font pRSM still exerted a unique and positive influence on errors ($\beta = 0.29$, p<0.005) indicating that gridfont letter pairs that are similar in a typical font are more likely to produce errors than letter pairs that are not. While the model accounted for the potential influence of the following pRSMs: stimulus-shape similarity, symbolic letter identity, phonetic letter-name similarity and motoric production codes—none of these reached significance (see Table 5 for details).

---

[6]While the magnitude of these R^2 measures may be small, it is worth noting that the LMEM has the ambitious goal of predicting trial-to-trial RTs (Model 1) or accuracies (Model 2). When a regression analysis is performed where the 5 pRSMs are used to predict the group Upright Gridfont RSM instead of individual trial-to-trial responses, the R^2 values increase substantially (Model 1 adjusted $R^2 = 0.55$, Model 2 adjusted $R^2 = 0.24$). This increase in model fit is likely due to the fact that trial and participant specific noise that is averaged out when combined into a group RSM.

Summary

Experiment 2 was designed to determine if findings from a different paradigm would converge with the findings from Experiment 1, while addressing some concerns about the open-endedness of the visual similarity judgment task used in Experiment 1. In fact, results obtained with the same/different task were remarkably similar to those observed in Experiment 1, with the pRSMS for Typical Font, Letter Identity, and Rotated Gridfont significantly predicting RTs, and the Typical Font pRSM predicting errors. The finding of the significant influence of the Typical Font pRSM provides further support for the role of allograph representations in letter recognition.

General discussion

An important question in research on letter processing and reading is whether or not stored structural descriptions of the spatial/geometrical features of canonical letter shapes—referred to here as allographs—play a role in the identification of letters presented in familiar and unfamiliar fonts. The analogous question for object recognition research concerns the human ability to recognize object despite vast amounts of surface variation and, therefore, the findings of the research reported here regarding letter recognition have implications for theories of object recognition more generally.

As we have indicated earlier, while theories of letter processing and reading typically posit allograph representations, the empirical evidence specifically supporting their role has been extremely limited. In two experiments using two different tasks—letter similarity judgments and same/different judgments—we combined Representational Similarity Analysis (RSA) with Linear Mixed Effects Modeling (LMEM) to identify the influence of different representational types on letter processing in these tasks. We used a novel empirical approach to address the challenge of distinguishing the contribution of computed stimulus-shape representations from allograph representations. We did so by using a gridfont that created differences between the similarity structures for these two representational types. We then indexed computed stimulus-shape similarity based on participant responses to rotated (unrecognizable) gridfont letters, while allograph similarity was indexed based on participant responses to upright letters presented in a typical font. Finally, in the data analysis, we examined the extent to which the similarity structure of three dependent variables (similarity judgments, same/different RTs and accuracy) could be explained by stimulus-shape, allograph, symbolic identity, phonetic and motor representations. We found that the only variable that consistently explained unique and significant variance for all three dependent variables, was allograph similarity.

Thus, the findings of the two experiments support the conclusion that when individuals process and recognize letters in a novel and unusual font, their responses are skewed/biased towards the visual similarity structure of more typical letter forms, supporting the hypothesis that we access stored allograph representations of the typical forms of letters during letter perception. These results provide evidence for cognitive models of reading that posit font-invariant allograph representation and provide constraints on models of reading without allograph representations in which, instead, the computed stimulus-shapes directly access either symbolic letter identities or lexical entries.

SLIs (symbolic letter identities)

The Letter Identity variable representing case-invariant letter identity (A=a) was shown to uniquely influence responses in both Experiments 1 and 2. By including visual (computed and stored), phonetic, and motoric similarity within the regression analyses, we were able to demonstrate—as would be predicted for purely symbolic representations—that this letter identity effect cannot be reduced to modality-specific visual, phonological or motor effects. This finding is also consistent with previous behavioral experimental evidence that cross-case letter forms share a common representation that is not modality-specific (e.g., Kinoshita & Kaplan, 2008; Wiley, Wilson, & Rapp, 2016). Consistent with the behavioral evidence of SLIs, Rothlein and Rapp (2014) provided neural data from an RSA analysis of fMRI data obtained from a letter decision task. They identified a region of left fusiform gyrus (in the vicinity of the VWFA, Cohen et al., 2000) that exhibited similar multi-voxel activation patterns in response to letters that had the same identity but different case. Importantly, this result could not be accounted for by visual, motor or phonetic similarity. The findings from the research reported here contribute convergent evidence from a novel paradigm that support the role of abstract symbolic letter representations, providing further constraints on theories of letter recognition and reading.

The finding of evidence for SLIs in this study is also relevant in addressing a potential concern that visual similarity judgments and same-different tasks might be based on task-specific representations and processes that do not inform our understanding of letter identification itself. However, it is worth noting that accessing stored letter representations such as SLIs in these tasks is not only extraneous, but in fact detrimental to the accuracy and response times in these tasks that only strictly require processing of visual features. Therefore, one is forced to ask—why are stored letter representations (e.g., allographs and SLIs) accessed in a task that does not require them? A likely explanation is that viewing letters (even in these tasks) automatically activates the representations involved in normal letter identification.

The representation of case (and font)

The results presented in this paper provide evidence for models of letter identification and reading in which allograph representations mediate between the computed stimulus shape representations and case-invariant SLIs. An important issue that these results do not address is how letter-case is represented. For example, in English orthography, letter-case representation is critical for differentiating nouns from proper nouns (bill and Bill), words from acronyms (pet and PET) and parsing sentences. Clearly, a complete theory of reading must include the representation of letter case (Grainger et al., 2016, Schubert & McCloskey, 2013). One possibility is that case information used in reading is available only (implicitly) either at the level of font-specific, computed letter shapes or at the level of allographs (Grainger et al., 2016; Brunsdon et al., 2006). In other words, case is represented by virtue of the fact that computed letter shapes and allographs **a**, *a*, and **A** are distinct from one another. Another possibility is that case is represented in terms of case-specific SLIs, such that **a** and *a* correspond to the same symbolic case-specific representation [a] while **A** corresponds to the different symbolic case-specific representation [A]. A third possibility is that case is represented as an independent abstract feature at the level of SLIs, such that the

SLI [A] could be associated either with an upper or lower case feature (Schubert & McCloskey, 2013). Note that these possibilities are not mutually exclusive.

In favor of the position that case is represented as a feature distinct and dissociable from letter identity, Schubert and McCloskey (2013) presented evidence from an individual (LHD) who acquired dyslexia subsequent to a stroke. The individual was administered a delayed-copy task in which a word was presented visually, covered from view and then she was asked to write it. In this and other visual letter processing tasks, the participant frequently perseverated letter identities from earlier responses to later ones. Importantly, in the delayed-copy task when words were presented in mixed case, the perseverated letter identities would assume the case of the letters they were substituting. For example, if the **R** in c-A-R perseverated into the word M-a-t it might be written as M-a-r but if it perseverated into the word M-a-T it would be produced as M-a-R. That the identity of the [R] in c-A-R could perseverate while assuming the case of the of the **t** in M-a-t is striking evidence that case is represented independently of letter identity.

Finally, although the focus in this study has been on font-independent representations, it is worth noting that font information, while not affecting the lexical identity of a word, can affect how a word is interpreted. Much in the same way that prosody can alter how a spoken word is interpreted, font (and case) can provide important emotional inflection to the meaning of a written word. For example, a word in italics can indicate sarcasm and a word in all uppercase letters can indicate excitement or shouting. Clearly, the representation of case and font in reading remain interesting and important research topics.

Levels of abstraction

In this research we have focused on three types/levels of representation that are claimed to play key roles in letter perception and reading: computed stimulus-shapes, allographs, and symbolic letter identities. Consistent with various specific proposals of letter perception and reading (e.g., Brunsdon et al., 2006; Dehaene et al., 2005; Grainger et al., 2008), these correspond to increasingly abstract levels of representations. Specifically, we have argued that computed stimulus-shape representations are font and case-specific, allographs are font invariant but case specific, while symbolic letter identities are both font and case-invariant (see Figure 1). The experimental paradigm and the RSA-LMEM analysis were designed to determine if there are unique contributions made to letter recognition by each of these hypothesized representational types. This is challenging because letter representations that are similar at one level, tend to be similar at other levels. However, the experimental paradigm and analysis approach were successful in allowing for this "cognitive dissection", revealing the independent influence of these representational types on letter perception.

As we have indicated, various theories of reading and letter recognition have posited a processing stream involving increasingly complex and abstract representations. While we agree with this general characterization, we would argue that this is not a "smooth" progression through representational abstraction. Instead, we see two distinct types of abstraction: *within-modality abstraction* and *amodal or symbolic abstraction*. The progression from increasingly more complex visual representations up through font-invariant allographs (structural descriptions) involves within-modality abstractions.

Representations and processes "morph" gradually from low-level visual representations to high-level spatial/geometric ones that can be considered to occur within the visuo-spatial modality. For letters, this maps onto the progression from computed-stimulus shapes to allographs. However, SLIs—because they are font and case-invariant—are not sensitive to visuo-spatial similarity. In fact, they represent visuo-spatially similar cross-case letter pairs (e.g., O/o, T/t, etc.) in the same way as dissimilar ones (e.g., A/a, G/g). Given their symbolic nature, the claim (see Figure 1) is that SLIs serve to translate between modalities. In fact, it is possible that it is in this respect that letters may be different from other objects as it is not clear that visual objects are represented in a comparable amodal, symbolic manner. This is an interesting question that is, however, outside the scope of the current paper.

Given the debate between abstractionist and exemplar/instance/grounded approaches to letter perception in reading (and a range of cognitive processes) it is worth considering whether or not the results reported here can be explained by exemplar/instance based accounts. As we indicated in the Introduction, recent computer programs of handwritten letter identification/classification have been highly successful using deep learning approaches that do not explicitly encode allographs or SLIs and can be considered to be exemplar or instance based. However, these models do involve multiple processing stages and—due to the nature of the learning and processing algorithms—the representational content of these stages is opaque. As a result, we simply do not know the types of representations that are involved and whether and to what extent they resemble the types of representations we have posited here.

Conclusions

In this work we deployed a novel paradigm and analysis approach to examine the cognitive representations involved in letter identification. This work provides evidence in favor of the hypothesis that allographs (structural descriptions) play a role in letter perception. In addition to adding to our understanding of letter perception and reading, the findings provide constraints on neuroimaging and computational work on these topics.

Acknowledgments

References

Baayen RH, Milin P. Analyzing reaction times. International Journal of Psychological Research. 2010; 3(2):12–28. http://doi.org/10.1287/mksc.12.4.395.

Barsalou LW. Grounded cognition. Annual Review of Psychology. 2008; 59:617–645. http://doi.org/10.1146/annurev.psych.59.103006.093639.

Barton, K. MuMIn: Multi-model inference. R package version 1.15.6. 2016. https://CRAN.R-project.org/package=MuMIn

Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models using lme4. Journal of Statistical Software. 2015; 67(1):1–48.

Besner D, Coltheart M, Davelaar E. Basic processes in reading: Computation of abstract letter identities. Canadian Journal of Psychology. 1984; 38(1):126–134. [PubMed: 6713295]

Boles DB, Clifford JE. An upper- and lowercase alphabetic similarity matrix, with derived generation similarity values. Behavior Research Methods, Instruments, & Computers. 1989; 21(6):579–586. http://doi.org/10.3758/BF03210580.

Bowers JS. In defense of abstractionist theories of repetition priming and word identification. Psychonomic Bulletin & Review. 2000; 7(1):83–99. [PubMed: 10780021]

Brunsdon R, Coltheart M, Nickels L. Severe developmental letter-processing impairment: A treatment case study. Cognitive Neuropsychology. 2006; 23(6):795–821. http://doi.org/10.1080/02643290500310863. [PubMed: 21049354]

Caplan LR, Hedley-Whyte T. Cuing and memory dysfunction in alexia without agraphia. Brain. 1974; 97:251–262. [PubMed: 4434176]

Carlson LA, Alejano AR, Carr TH. The level-of-focal-attention hypothesis in oral reading: Influence of strategies on the context specificity of lexical repetition effects. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1991; 17(5):924–931. http://doi.org/10.1037/0278-7393.17.5.924.

Chanoine V, Ferreira CT, Demonet JF, Nespoulous JL, Poncet M. Optic aphasia with pure alexia: a mild form of visual associative agnosia? A case study. Cortex. 1998; 34(3):437–448. http://doi.org/10.1016/S0010-9452(08)70766-4. [PubMed: 9669108]

Chauncey K, Holcomb PJ, Grainger J. Effects of stimulus font and size on masked repetition priming: An event-related potentials (ERP) investigation. Language and Cognitive Processes. 2008; 23(1):183–200. http://doi.org/10.1080/01690960701579839. [PubMed: 19590754]

Cohen L, Dehaene S, Naccache L, Lehéricy S, Dehaene-Lambertz G, Hénaff Ma, Michel F. The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. Brain. 2000; 123:291–307. [PubMed: 10648437]

Cox CH, Coueignoux P, Blesser B, Eden M. Skeletons: A link between theoretical and physical letter descriptions. Pattern Recognition. 1982; 15(1):11–22. http://doi.org/10.1016/0031-3203(82)90056-5.

Dalmás JF, Dansilio S. Visuographemic alexia: A new form of a peripheral acquired dyslexia. Brain and Language. 2000; 75(1):1–16. http://doi.org/10.1006/brln.2000.2321. [PubMed: 11023635]

Davitt LI, Cristino F, Wong AC-N, Leek EC. Shape information mediating basic- and subordinate-level object recognition revealed by analyses of eye movements. Journal of Experimental Psychology. Human Perception and Performance. 2014; 40(2):451–456. http://doi.org/10.1037/a0034983. [PubMed: 24364701]

Dehaene S, Cohen L, Sigman M, Vinckier F. The neural code for written words: a proposal. Trends in Cognitive Sciences. 2005; 9(7):335–341. http://doi.org/10.1016/j.tics.2005.05.004. [PubMed: 15951224]

Egeth H, Blecker D. Differential effects of familiarity on judgments of sameness and difference. Perception & Psychophysics. 1971; 9(4):321–326.

Gauthier I, Tarr MJJ, Moylan J, Skudlarski P, Gore JC, Anderson AW. The fusiform "face area" is part of a network that processes faces at the individual level. Journal of Cognitive Neuroscience. 2000; 12(3):495–504. [PubMed: 10931774]

Gauthier I, Wong AC-N, Hayward WG, Cheung OS. Font tuning associated with expertise in letter perception. Perception. 2006; 35(4):541–559. http://doi.org/10.1068/p5313. [PubMed: 16700295]

Goldinger SD. Echoes of echoes? An episodic theory of lexical access. Psychological Review. 1998; 105(2):251–279. [PubMed: 9577239]

Grainger J, Rey A, Dufau S. Letter perception: From pixels to pandemonium. Trends in Cognitive Sciences. 2008; 12(10):381–387. http://doi.org/10.1016/j.tics.2008.06.006. [PubMed: 18760658]

Grainger J, Dufau S, Ziegler JC. A vision of reading. Trends in Cognitive Sciences. 2016; 20(3):171–179. http://doi.org/10.1016/j.tics.2015.12.008. [PubMed: 26809725]

Grant SC, Logan GD. The loss of repetition priming and automaticity over time as a function of degree of initial learning. Memory & Cognition. 1993; 21(5):611–618. http://doi.org/10.3758/BF03197193. [PubMed: 8412713]

Graves A, Liwicki M, Fernandez S, Bertolami R, Bunke H, Schmidhuber J. A novel connectionist system for unconstrained handwriting recognition. IEEE Transactions on Pattern Analysis and

Machine Intelligence. 2009; 31(5):855–868. http://doi.org/10.1109/TPAMI.2008.137. [PubMed: 19299860]

Herrick EM. Letters with alternative basic shapes. Visible Language. 1975; 2(9):133–144.

Hofstadter, DR., McGraw, G. Fluid Concepts and Creative Analogies. Computer models of the fundamental mechanisms of thought. New York, NY: BasicBooks; 1995. Letter Spirit: Esthetic perception and creative play in the rich microcosm of the Roman alphabet; p. 407-466.

Kinoshita S, Kaplan L. Priming of abstract letter identities in the letter match task. Quarterly Journal of Experimental Psychology (2006). 2008; 61(12):1873–1885. http://doi.org/10.1080/17470210701781114. [PubMed: 18609367]

Konkle T, Brady TF, Alvarez GA, Oliva A. Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. Journal of Experimental Psychology: General. 2010; 139(3): 558–578. http://doi.org/10.1037/a0019165. [PubMed: 20677899]

Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis - connecting the branches of systems neuroscience. Frontiers in Systems Neuroscience. 2008; 2(4):1–28. http://doi.org/10.3389/neuro.06.004.2008. [PubMed: 18958245]

Kuznetsova, A., Brockhoff, PB., Christensen, RHB. lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-32. 2016. https://CRAN.R-project.org/package=lmerTest

Lupyan G, Thompson-Schill SL, Swingley D. Conceptual penetration of visual processing. Psychological Science. 2010; 21(5):682–691. http://doi.org/10.1177/0956797610366099. [PubMed: 20483847]

Marr D, Nishihara HK. Representation and recognition of the spatial organization of three-dimensional shapes. Proceedings of the Royal Society of London. 1978; 200(1140):269–294. http://doi.org/10.1098/rspb.1978.0020. [PubMed: 24223]

Marsolek C. Abstractionist versus exemplar-based theories of visual word priming: a subsystems resolution. The Quarterly Journal of Experimental Psychology. 2004; 57(7):1233–1259. http://doi.org/10.1080/02724980343000747. [PubMed: 15513245]

Miozzo M, Caramazza A. Varieties of pure alexia: The case of failure to access graphemic representations. Cognitive Neuropsychology. 1998; 15:203–238.

Mueller ST, Weidemann CT. Alphabetic letter identification: Effects of perceivability, similarity, and bias. Acta Psychologica. 2012; 139(1):19–37. http://doi.org/10.1016/j.actpsy.2011.09.014. [PubMed: 22036587]

Naber M, Nakayama K. Pupil responses to high-level image content. J. ournal of Vision. 2013; 13(6): 1–8. 7, http://doi.org/10.1167/13.6.7.doi.

Petit J-P, Midgley KJ, Holcomb PJ, Grainger J. On the time course of letter perception: A masked priming ERP investigation. Psychonomic Bulletin & Review. 2006; 13(4):674–681. [PubMed: 17201369]

Qiao E, Vinckier F, Szwed M, Naccache L, Valabrègue R, Dehaene S, Cohen L. Unconsciously deciphering handwriting: Subliminal invariance for handwritten words in the visual word form area. NeuroImage. 2010; 49(2):1786–1799. http://doi.org/10.1016/j.neuroimage.2009.09.034. [PubMed: 19781648]

Rapp BC, Caramazza A. Letter processing in reading and spelling: Some dissociations. Reading and Writing: An Interdisciplinary Journal. 1989; 1:3–23.

Rapp B, Caramazza A. From graphemes to abstract letter shapes: Levels of representation in written spelling. Journal of Experimental Psychology: Human Perception and Performance. 1997; 23(4): 1130–1152. [PubMed: 9269731]

Rapp B, Lipka K. The literate brain: The relationship between spelling and reading. Journal of Cognitive Neuroscience. 2011; 23:1180–1197. [PubMed: 20433242]

Riesenhuber M, Poggio T. Hierarchical models of object recognition in cortex. Nature Neuroscience. 1999; 2(11):1019–1025. http://doi.org/10.1038/14819. [PubMed: 10526343]

Rothlein D, Rapp B. The similarity structure of distributed neural responses reveals the multiple representations of letters. NeuroImage. 2014; 89:331–344. http://doi.org/10.1016/j.neuroimage.2013.11.054. [PubMed: 24321558]

Sanocki T. Effects of font- and letter-specific experience on the perceptual processing of letters. The American Journal of Psychology. 1992; 105(3):435–458.

Sanocki T, Dyson MC. Letter processing and font information during reading: beyond distinctiveness, where vision meets design. Attention, Perception & Psychophysics. 2012; 74(1):132–145. http://doi.org/10.3758/s13414-011-0220-9.

Schubert T, McCloskey M. Prelexical representations and processes in reading: Evidence from acquired dyslexia. Cognitive Neuropsychology. 2013; 30(6):360–395. http://doi.org/10.1080/02643294.2014.880677. [PubMed: 24512594]

Shepard RN, Chipman S. Second-order isomorphism of internal representations: Shapes of states. Cognitive Psychology. 1970; 1:1–17.

Simpson IC, Mousikou P, Montoya JM, Defior S. A letter visual-similarity matrix for Latin-based alphabets. Behavior Research Methods. 2012; 45(2):431–439. http://doi.org/10.3758/s13428-012-0271-4.

Tenpenny PL. Abstractionist versus episodic theories of repetition priming and word identification. Psychonomic Bulletin & Review. 1995; 2(3):339–363. [PubMed: 24203715]

Tulving, E. Essentials of episodic memory. Press, OU., editor. New York: 1983.

Walker P. Font tuning: A review and new experimental evidence. Visual Cognition. 2008; 16(8):1022–1058. http://doi.org/10.1080/13506280701535924.

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

Watson MR, Akins KA, Enns JT. Second-order mappings in grapheme-color synesthesia. Psychonomic Bulletin & Review. 2012; 19(2):211–217. http://doi.org/10.3758/s13423-011-0208-4. [PubMed: 22231728]

Wiley RW, Wilson C, Rapp B. The effects of alphabet and expertise on letter perception. Journal of Experimental Psychology: Human Perception and Performance. 2016; 42(8):1186–1203. http://dx.doi.org/10.1037/xhp0000213. [PubMed: 26913778]

Wilson M. Six views of embodied cognition. Psychonomic Bulletin & Review. 2002; 9(4):625–636. [PubMed: 12613670]

Wong AC-N, Gauthier I. An analysis of letter expertise in a levels-of-categorization framework. Visual Cognition. 2007; 15(7):854–879. http://doi.org/10.1080/13506280600948350.
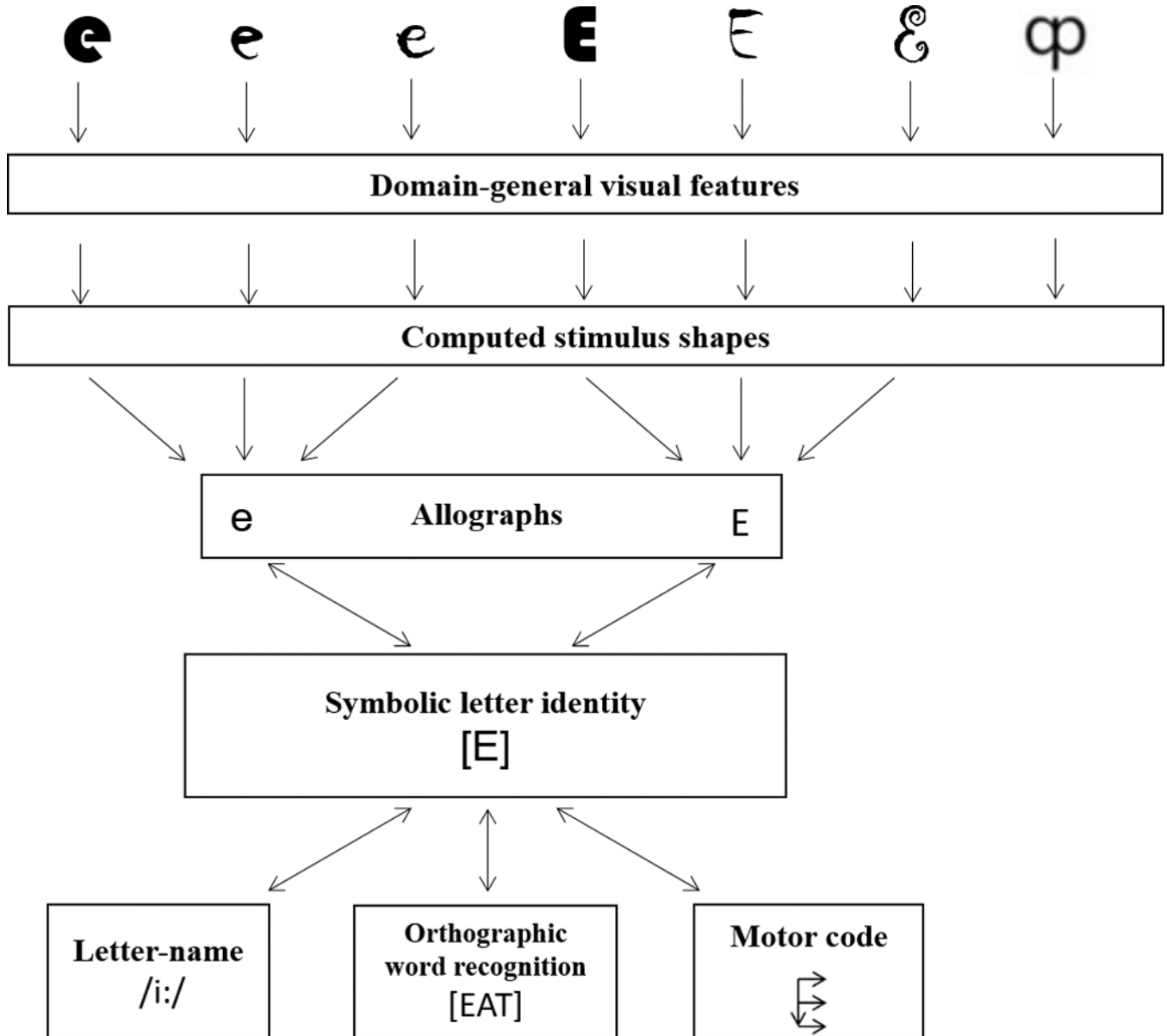
Figure 1.
Proposed representational levels in the processing of different fonts of upper and lower-case E and one pseudoletter. In the initial processing stages, stimuli are represented in terms of domain-general visual features (e.g. simple and complex cells representing oriented bars). Many models also posit a font-specific computed stimulus shape representations of the shape of the stimulus independent of its identity. This level of representation would allow us to describe the shape of the letter E as well as the shape of any given pseudoletter. Computed stimulus shape representations do not encode letter identity information or even whether the shape is a letter or not. These computed stimulus shape representations go on to access stored font-invariant allograph representations that presumably encode letter shapes in a manner that abstracts away from certain differences in stimulus font. The allographs in turn activate abstract or symbolic letter identity (SLI) representations that are both font and case

independent—abstracting away from visual information altogether. SLI representations serve as input to the lexical and sublexical reading processes that mediate orthographic word recognition. They also serve as a conduit to cross-modal letter representations like phonological letter names and motoric production codes for written letter production.
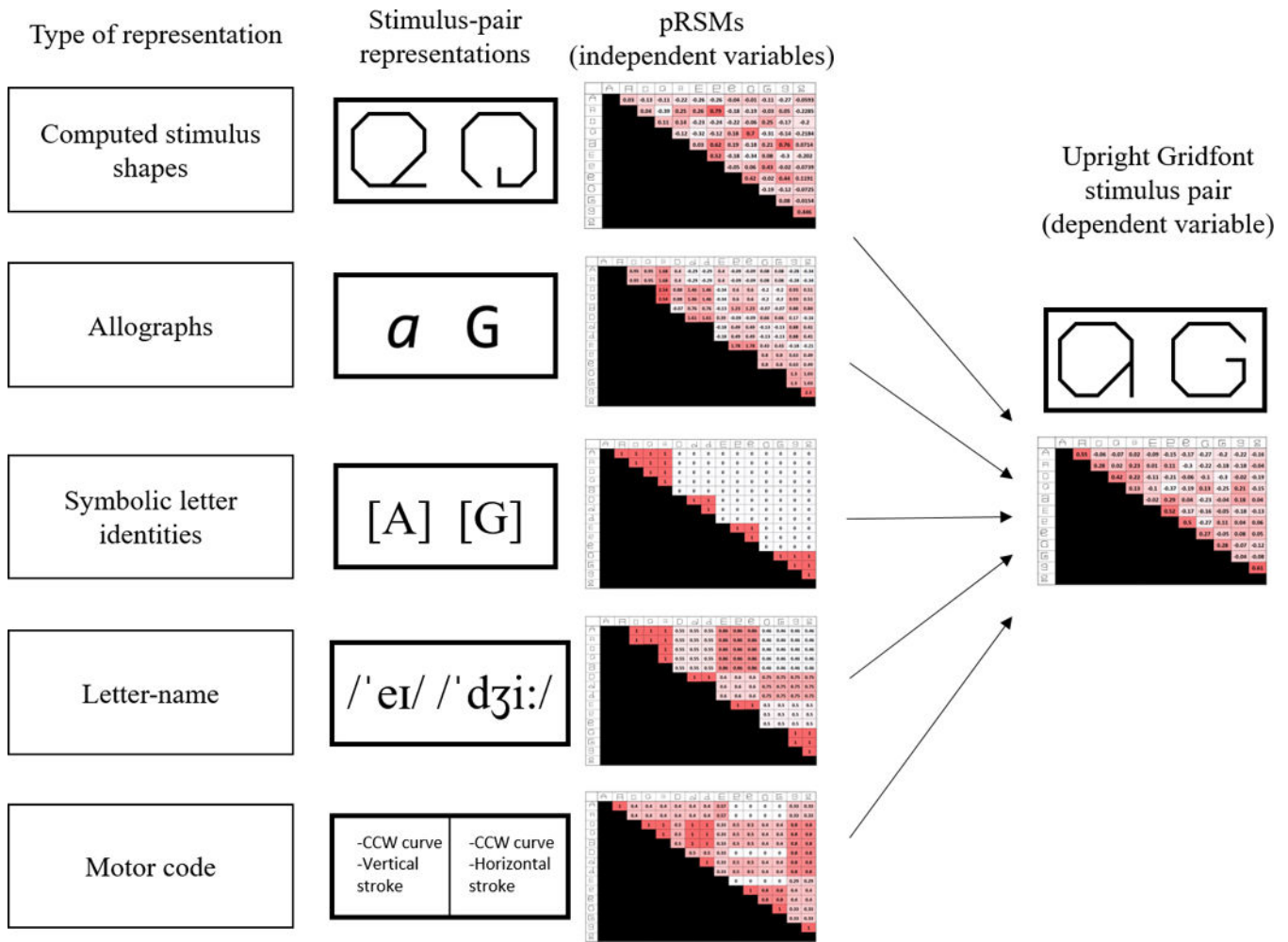
Figure 2.
Using Representational Similarity Analysis to identify the influence of specific representational types on the recognition of letters viewed in the Upright Gridfont. **(A)** The first column depicts the types of letter representations potentially active in response to viewing *a* and **G** in the Upright Gridfont. **(B)** The second column depicts an approximation of the representational content of Upright Gridfont *a* and **G** for each representational type. **Computed Stimulus-Shape** similarity is estimated with rotated Gridfont letters. Rotation maintains pairwise visual-spatial similarity while minimizing the identifiability of the rotated letters, thereby, limiting the influence of the other representational types on letter recognition. **Allograph** similarity is estimated from the similarity structure of a more typical font. **Symbolic Letter identities** encode font and case-invariant letter identity. **Letter-name** similarity is estimated from the phonetic features that compose the phonemes of the letter-names. **Motoric** similarity is estimated from hypothesized motor plans. **(C)** The third column depicts matrices of pairwise similarity estimates for each type of representation (pRSMs), characterizing the predicted similarity structures at each level of representation in response to pairs of the Upright Gridfont letters. **(D)** Column 4 depicts a set of similarity responses (e.g., derived from judgments or same/different RTs or accuracy) elicited in response to letter pairs (the oRSM). These similarity responses are tested for the unique
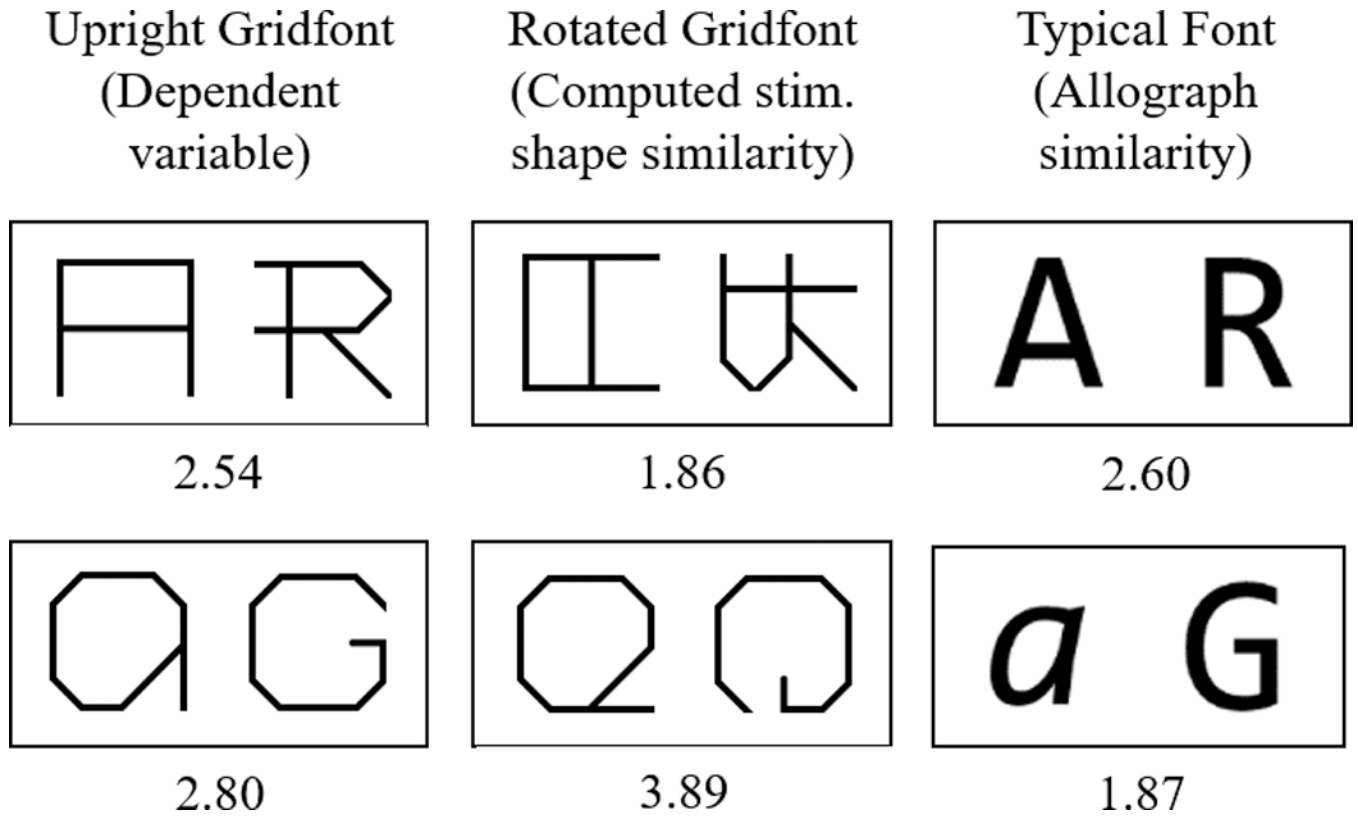
influence of each of the 5 types of letter representations by running a linear mixed-effects model (LMEM) in which the 5 pRSMs simultaneously predict the responses to the Upright Gridfont letter-pairs.

| Upright Gridfont (Dependent variable)	Rotated Gridfont (Computed stim. shape similarity)	Typical Font (Allograph similarity)
2.54	1.86	2.60
2.80	3.89	1.87

Figure 3.

Dissociating the influence of Computed Stimulus-Shape Representations from Allograph Representations. Participants made visual similarity judgments to letters presented in an atypical Upright Gridfont (left column), a Rotated Gridfont (middle column), and a Typical Font (right column). The average similarity judgment (computed from Experiment 1) is shown below each example letter-pair. A response of 1 indicates little or no visual similarity for the pair and 5 indicates the highest level of visual similarity. If computed stimulus-shape representations are the only type of letter representation that contributed to visual similarity judgments, we would predict similarity judgments for the Upright Gridfont (left column) and the Rotated Gridfont (middle column) to be identical. Instead, we see that the similarity judgments to the Upright Gridfont are different from similarity judgments to the Rotated Gridfont. The top row depicts an example where there is a bias to judge the Upright Gridfont letter-pair as being more similar than would be predicted by the computed stimulus-shape similarity alone while the bottom row shows the opposite effect. The fact that the Upright Gridfont similarity judgments are "pulled" in the direction of the Typical Font similarity judgments suggests an influence of allograph representations on these similarity judgments
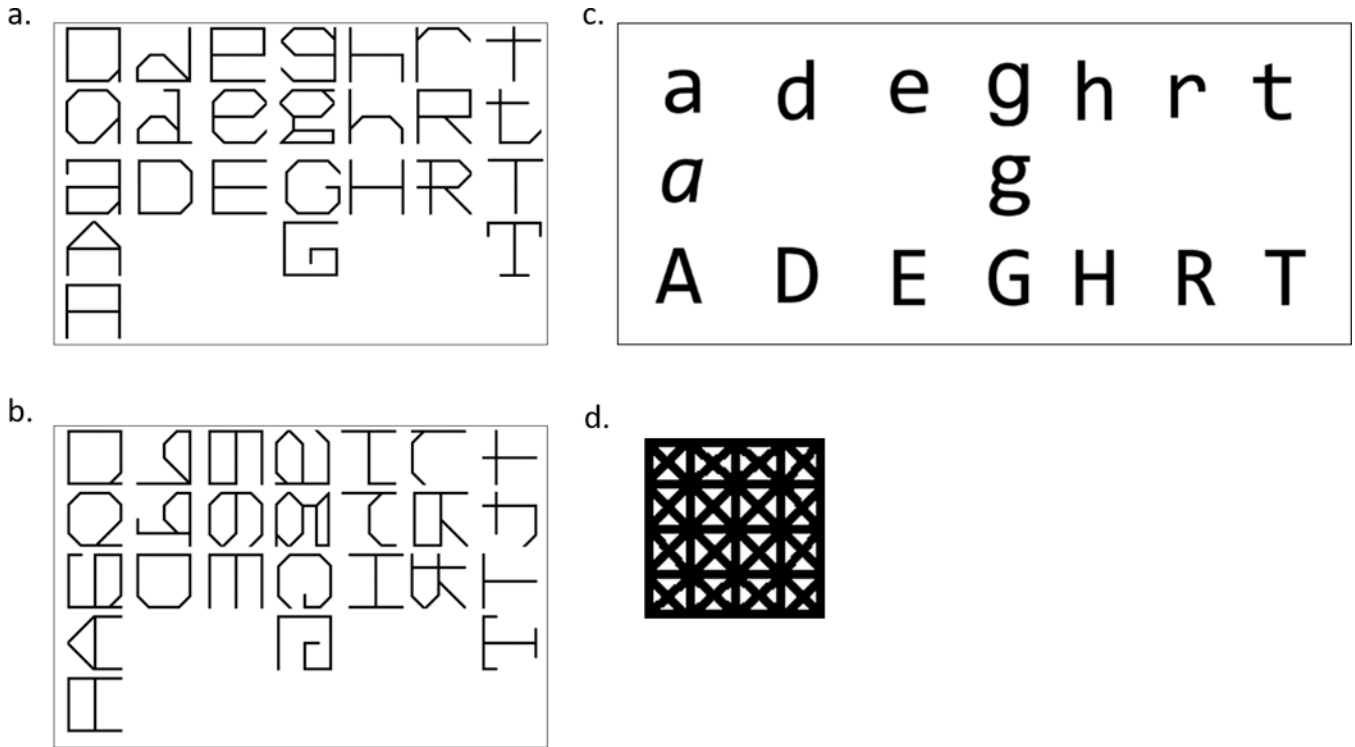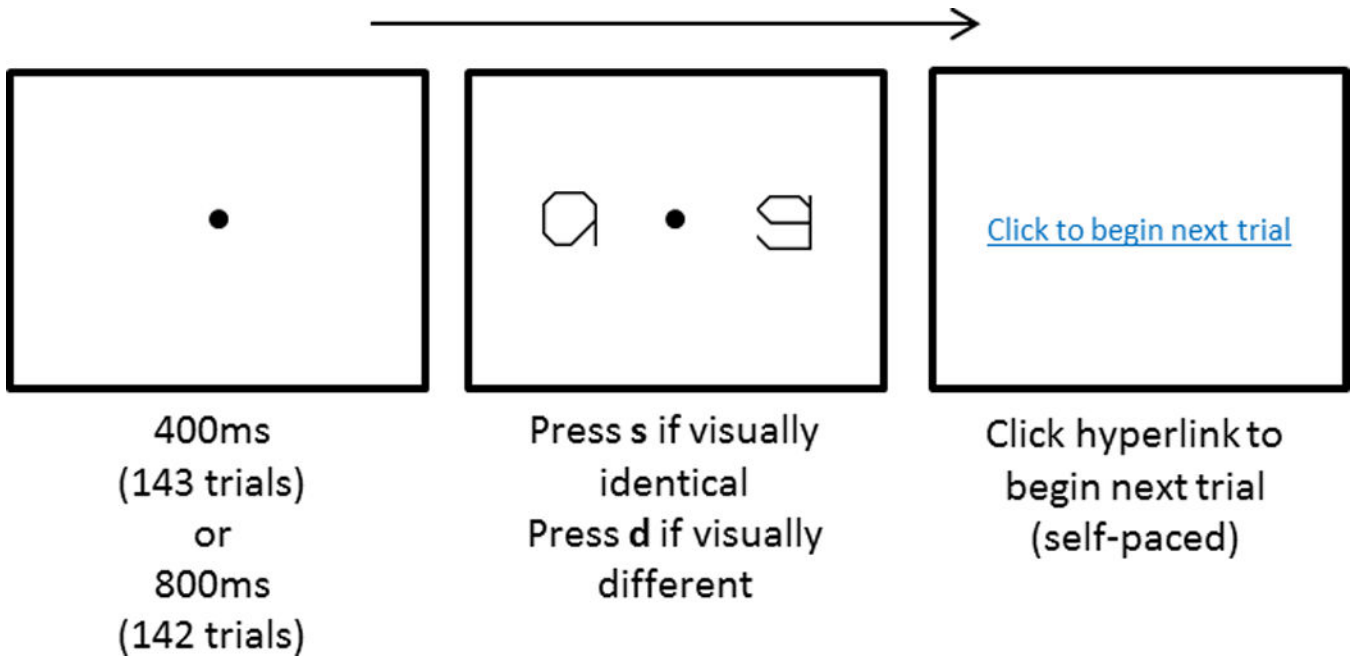
Figure 4.
Experimental stimuli. a) The novel, atypical and Upright Gridfont alphanumeric characters used as stimuli in Experiments 1. Experiment 2 used the same stimuli except the 6 stimuli corresponding to the identities [D] or [H]. b) The Rotated Gridfont stimuli used in Experiments 1 and 2. These stimuli were created by flipping the upright characters about their vertical axis and then rotating them 90°ccw. The rotation intended to render them difficult to identify. Experiment 2 did not include the 6 stimuli corresponding to rotated versions of [D] or [H]. c) Font used in the Typical Font Group. The font is Consalas which is a variant of Calibri where the width of each character is matched. d) A depiction of the square, 100 X 100 pixel grid with all of the possible features in black. Each stimulus was required to touch all 4 sides, ensuring the height and width of each stimulus were matched.
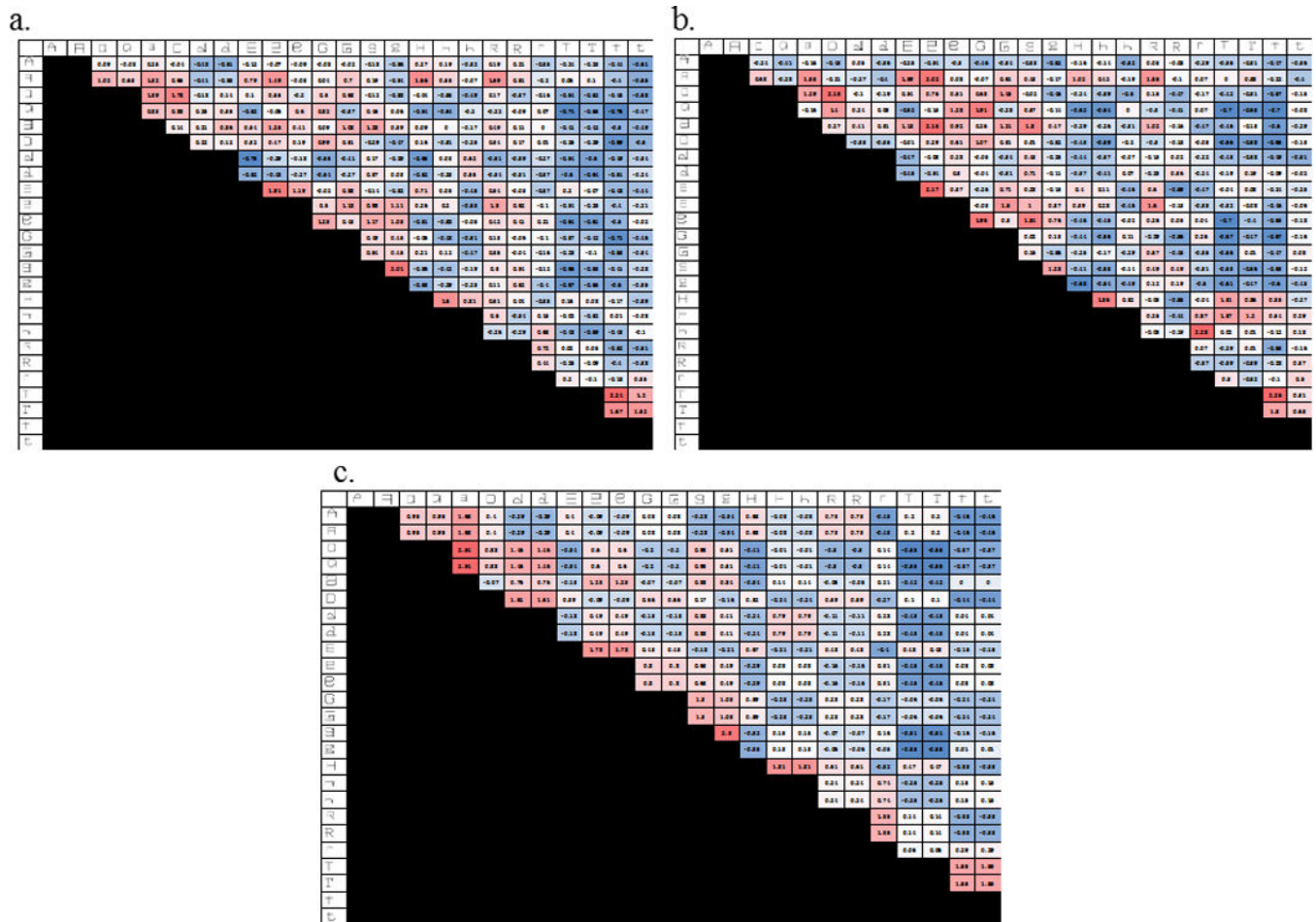
Figure 5.
The group representational similarity matrices (RSMs) obtained from Experiment 1. Red cells indicate larger average similarity judgments and blue cells indicate smaller average similarity judgments. a) The group RSM for the similarity judgments to letter-pairs presented in the atypical Upright Gridfont. The Gridfont letters are depicted on the margin. The group RSM was formed by normalizing each participant's responses to have a mean of 0 and a standard deviation of 1 and then averaging the normalized responses to each letter-pair across participants. b) The group RSM for the similarity judgments to letter-pairs presented in the atypical *Rotated* Gridfont. c) The group RSM for the similarity judgments to letter-pairs presented in the Typical font.

Figure 6.
Example trial in the same-different decision paradigm.

a.

	A	A	□	C	a	E	P	e	G	G	9	8	R	R	r	T	T	+	t
A		0.55	-0.06	-0.07	0.02	-0.09	-0.15	-0.17	-0.27	-0.2	-0.22	-0.16	0.01	-0.11	-0.05	-0.3	-0.15	-0.11	0.01
A			0.28	0.02	0.23	0.01	0.11	-0.3	-0.22	-0.18	-0.18	-0.04	0.28	-0.37	-0.13	-0.06	-0.24	-0.29	-0.22
□				0.42	0.22	-0.11	-0.21	-0.06	-0.1	-0.3	-0.02	-0.19	-0.17	-0.28	-0.14	-0.2	-0.03	-0.28	-0.06
C					0.13	-0.1	-0.37	-0.19	0.13	-0.25	0.21	-0.15	-0.1	0.03	-0.09	-0.23	-0.27	0.01	-0.02
a						-0.02	0.29	0.04	-0.23	-0.04	0.18	0.04	-0.09	-0.09	-0.15	-0.25	-0.09	-0.32	-0.28
E							0.52	-0.17	-0.16	-0.05	-0.18	-0.13	-0.12	-0.4	-0.29	-0.25	-0.07	-0.43	-0.36
P								0.5	-0.27	0.11	0.04	0.06	0.04	-0.16	-0.26	-0.15	-0.2	-0.24	-0.21
e									0.27	-0.05	0.08	0.05	-0.1	-0.3	-0.11	-0.28	-0.26	-0.39	-0.3
G										0.28	-0.07	-0.12	-0.31	-0.24	-0.1	-0.43	-0.27	-0.39	-0.22
G											-0.04	-0.08	-0.01	-0.24	-0.11	-0.18	-0.17	-0.21	-0.31
9												0.61	-0.17	-0.2	-0.26	-0.08	-0.26	-0.31	-0.3
8													0	0.12	-0.21	-0.34	-0.24	-0.32	-0.24
R														0.89	-0.17	-0.32	-0.38	-0.31	-0.33
R															-0.11	-0.27	-0.38	-0.28	-0.15
r																0.01	-0.32	-0.17	-0.2
T																	0.72	0.62	0.05
T																		0.16	-0.01
+																			0.36
t																			

b.

	A	A	□	a	B	E	P	e	G	G	9	8	R	R	r	T	T	+	t
A		0.03	-0.13	-0.11	-0.22	-0.26	-0.26	-0.04	-0.01	-0.11	-0.27	-0.06	0.14	-0.03	-0.14	-0.01	-0.16	-0.21	-0.07
A			0.04	-0.39	0.25	0.26	0.79	-0.18	-0.19	-0.03	0.05	-0.23	0.51	-0.14	-0.39	-0.07	0.18	-0.2	-0.19
□				0.11	0.14	-0.23	-0.24	-0.22	-0.06	0.25	-0.17	-0.2	-0.06	-0.28	-0.12	-0.26	-0.25	-0.21	-0.44
a					-0.12	-0.32	-0.12	0.18	0.7	-0.31	-0.14	-0.22	-0.14	-0.26	-0.24	-0.39	-0.27	-0.39	-0.31
B						0.03	0.62	0.19	-0.18	0.21	0.76	0.07	0.08	-0.38	0.02	-0.46	-0.06	-0.34	-0.23
E							0.52	-0.18	-0.34	0.08	-0.3	-0.2	-0.08	-0.35	-0.14	-0.08	0	-0.24	-0.19
P								-0.05	0.06	0.43	-0.02	-0.07	0.32	-0.28	-0.35	-0.13	-0.02	-0.35	-0.12
e									0.42	-0.02	0.44	0.12	-0.21	-0.13	0.01	-0.31	-0.24	-0.28	-0.32
G										-0.19	-0.12	-0.07	-0.24	-0.23	-0.26	-0.22	-0.2	-0.21	-0.33
G											0.08	-0.02	0.18	-0.26	-0.17	-0.09	-0.15	-0.12	-0.17
9												0.45	0.25	0.13	-0.31	-0.37	-0.09	-0.36	-0.22
8													-0.07	0.1	-0.31	-0.24	-0.17	-0.35	-0.33
R														0.15	-0.09	-0.09	-0.07	-0.14	-0.25
R															-0.07	-0.2	-0.14	-0.16	0.02
r																-0.09	-0.07	-0.24	-0.03
T																	0.1	1.14	0.06
T																		-0.02	0.03
+																			-0.22
t																			

Figure 7.
The group representational similarity matrices (RSMs) computed from the RTs to different trials in Experiment 2. Red cells indicate slower average RTs (more similarity) and blue cells indicate faster RTs (less similarity). a) The group RSM for the RTs to letter-pairs presented in the Upright Gridfont. The Gridfont letters are depicted on the margin. The group RSM was formed by normalizing each participant's responses to have a mean of 0 and a standard deviation of 1 and then averaging the normalized responses to each letter-pair across

participants. b) The group RSM for the average normalized RTs to letter-pairs presented in the Rotated Gridfont.

a.

b.

Figure 8.
The group representational similarity matrices (RSMs) depicting the error counts to different trials across participants in Experiment 2. *Errors were defined as "same" responses to "different" trials.* Darker cells indicate a greater error count. a) The group RSM for the RTs to letter-pairs presented in the Upright Gridfont. The Gridfont letters are depicted on the margin. The group RSM was formed by normalizing each participant's responses to have a mean of 0 and a standard deviation of 1 and then averaging the normalized responses to each

letter-pair across participants. b) The group RSM for the average normalized RTs to letter-pairs presented in the Rotated Gridfont.

Table 1

Cross-correlations of each of the RSMs used in or developed from Experiment 1. The Upright gridfont matrix serves as the dependent measure (oRSM) while the others serve as predictor variables (pRSMs) See text for details on how each matrix was derived.

	Experiment 1 (Similarity Judgments) – RSM correlations					
	Upright Gridfont	Rotated Gridfont	Typical Font	Letter Identity	Phonetic Similarity	Motor Similarity
Upright Gridfont	-	0.78	0.69	0.47	0.06	0.17
Rotated Gridfont	-	-	0.40	0.20	0.03	0.13
Typical Font	-	-	-	0.68	0.29	0.31
Letter Identity	-	-	-	-	0.49	0.20
Phonetic Similarity	-	-	-	-	-	0.15
Motor Similarity	-	-	-	-	-	-

Note. RSM = Representational Similarity Matrix

Table 2

Results from the LMEM-RSA analyses for visual similarity judgments of pairs of Upright Gridfont letters from Experiment 1. Betas are standardized against the other fixed effect predictors.

| | Experiment 1 (Similarity Judgments) | | | |
	beta	SE	df	*t* value	p
Typical Font pRSM	0.21	0.03	126	7.12	<0.0001
Letter Identity pRSM	0.08	0.03	94	2.19	0.03
Rotated Gridfont pRSM	0.34	0.03	79	11.49	<0.0001
Phonetic Similarity pRSM	−0.09	0.02	173	−3.85	0.0002
Motor Similarity pRSM	−0.03	0.02	188	−1.36	0.17
Trial Order	0.00	0.01	12880	−0.34	0.73
Response RT	0.34	0.02	12990	19.16	<0.0001

Note. df = degrees of freedom; SE = standard error; pRSM = predicted Representational Similarity Matrix

Table 3

Cross-correlations of each of the RSMs used in or developed from the RT analysis (Model 2) in Experiment 2. The average normalized RTs across participants to the Upright Gridfont composed the Upright Gridfont RSM and served as the dependent measure (oRSM) while the other RSMs served as predictor variables (pRSMs). The Rotated Gridfont consisted of the average normalized RTs across participants who saw the Rotated Gridfont. See text for details on how the other RSMs were derived.

| | Experiment 2 (RTs) – RSM correlations | | | | |
	Upright Gridfont	Rotated Gridfont	Typical Font	Letter Identity	Phonetic Similarity	Motor Similarity
Upright Gridfont	-	0.56	0.56	0.46	0.23	0.29
Rotated Gridfont	-	-	0.42	0.21	0.12	0.19
Typical Font	-	-	-	0.68	0.29	0.37
Letter Identity	-	-	-	-	0.49	0.13
Phonetic Similarity	-	-	-	-	-	0.06
Motor Similarity	-	-	-	-	-	-

Note. RSM = Representational Similarity Matrix

Table 4

Cross-correlations of each of the RSMs used in or developed from the error analysis (Model 2) in Experiment 2. The error-count ("same" responses to "different" trials) across participants from the Upright Gridfont matrix served as the dependent measure (oRSM) while the others served as predictor variables (pRSMs). The Rotated Gridfont consists of the error-count across participants who saw the Rotated Gridfont. See text for details on how the other RSMs were derived.

| | Experiment 2 (Errors) – RSM correlations | | | | | |
	Upright Gridfont	Rotated Gridfont	Typical Font	Letter Identity	Phonetic Similarity	Motor Similarity
Upright Gridfont	-	0.12	0.47	0.42	0.21	0.27
Rotated Gridfont	-	-	0.16	0.18	0.15	0.28
Typical Font	-	-	-	0.68	0.29	0.31
Letter Identity	-	-	-	-	0.49	0.20
Phonetic Similarity	-	-	-	-	-	0.15
Motor Similarity	-	-	-	-	-	-

Note. RSM = Representational Similarity Matrix

Table 5

Results from the LMEM-RSA analyses for RTs (model 1) and accuracy (model 2) for a same/different task (Experiment 2). Betas are standardized against the other fixed effect predictors.

	Experiment 2 – Model 1 (Different RTs)					Experiment 2 – Model 2 (Same/Diff Accuracy)			
	beta	SE	DF	t value	p	beta	SE	Z value	p
Typical Font pRSM	0.011	0.005	147	2.30	0.02	0.29	0.10	2.95	0.003
Letter Identity pRSM	0.012	0.004	160	2.78	0.01	0.10	0.10	1.05	0.30
Rotated Gridfont pRSM	0.025	0.003	163	7.52	<0.0001	0.05	0.05	0.99	0.32
Phonetic Similarity pRSM	−0.004	0.004	157	−0.91	0.37	−0.09	0.10	−0.84	0.40
Motor Similarity pRSM	−0.004	0.004	162	−1.05	0.29	0.03	0.09	0.40	0.69
Trial Order	0.011	0.002	7079	4.73	<0.0001	−0.06	0.07	−0.79	0.43
Prev. Trial RT	0.13	0.009	7129	14.34	<0.0001				

Note. df = degrees of freedom; SE = standard error; pRSM = predicted Representational Similarity Matrix; RT = reaction time