Letter string processing and visual short-term memory

Maria Ktori a, Jonathan Grainger a & Stéphane Dufau a
a CNRS and Aix-Marseille University, Marseille, France

Available online: 19 Aug 2011
Letter string processing and visual short-term memory

Maria Ktori, Jonathan Grainger, and Stéphane Dufau
CNRS and Aix-Marseille University, Marseille, France

The present study investigated whether expertise with letter string processing influences visual short-term memory capacity. Specifically, we examined whether performance in a change-detection task would vary as a function of stimulus type (letters vs. symbols) and type of display (horizontal, vertical, and circular). Participants were asked to detect a one-character change in a briefly presented character array following a delay of 900 ms. Concurrent articulation was used to limit effects of rehearsal. Type of display significantly affected performance with letters, but not with symbols, with a selective increase in change-detection accuracy for horizontally presented letter arrays compared with vertical and circular arrays. These findings confirm the standard limits of storage in visual short-term memory, but critically reveal a selective advantage for letter arrays over symbol arrays when presented horizontally. Such an advantage is probably due to the utilization of a specialized encoding mechanism built up over years of reading experience.

**Keywords:** Letter string processing; Visual short-term memory; Expertise.

Visually processed information can be maintained in memory for a short period of time, to be either subsequently acted upon if needed, or discarded if not. The system that maintains this information is known as visual short-term memory (VSTM, hereafter). VSTM is acknowledged as an important component of perceptual and cognitive processing in tasks that depend on visual input (e.g., Jolicœur & Dell'Acqua, 1998; Jolicœur, Sessa, Dell'Acqua, & Robitaille, 2006a, 2006b; Prime & Jolicœur, 2010). VSTM representations are created rapidly (Gegenfurtner & Sperling, 1993; Shibuya & Bundesen, 1988; Vogel, Woodman, & Luck, 2006) and are coarsely specified (Simons & Rensink, 2005), and information is thought to be retained by an active mechanism that is implemented by sustained neural firing (Luck & Hollingworth, 2008). A distinctive characteristic of VSTM is its limited storage capacity, which is estimated to be around four simple objects (Cowan, 2001; Luck & Vogel, 1997; Phillips, 1974).

However, there is evidence to suggest some kind of flexibility in the limits of VSTM capacity as the result of the processing strategy used to encode objects. This evidence comes particularly from studies of perceptual expertise. Perceptual expertise arises from extensive experience with items within a
specific domain and refers to the enhanced ability to make subordinate-level discriminations between them (Scolari, Vogel, & Awh, 2008). Furthermore, this ability can be transferred to new items within the trained domain (Curby & Gauthier, 2010). As such, visual experts process highly complex objects with relative ease if these belong within their domain of expertise.

Previous research exploring whether expert perception impacts VSTM has focused mainly on the differences between face processing and the processing of ordinary objects (Curby & Gauthier, 2007; Curby, Glazek, & Gauthier, 2009; Scolari et al., 2008). This is because faces represent a class of stimuli of perceptual expertise to the typical observer, and, as such, discrimination of upright faces is one of the best documented examples of expert perception (Scolari et al., 2008). However, there is also research that has explored the relationship between expert perception and VSTM capacity in other domains of expertise such as chess playing (Chase & Simon, 1973; Freyhof, Gruber, & Ziegler, 1992; Gobet & Simon, 1998) and car knowledge (Curby et al., 2009). Findings from these studies provide supportive evidence of an enhancing effect of perceptual expertise on VSTM. VSTM capacity for domain-specific items is greater in perceptual experts than in nonexpert individuals (see Curby & Gauthier, 2010, for a review and relevant discussions). This effect is attributed to the holistic (or configural) processing associated with face processing and expert perception in general. It is believed that this specialized process takes into account information about individual features as well as the relationship between them (Curby & Gauthier, 2010; Farah, Wilson, Drain, & Tanaka, 1998; Scolari et al., 2008; Tanaka & Sengco, 1997). The utilization of such a process enables an encoding mechanism that may incorporate more and/or tightly bound features into the unified VSTM representations (Scolari et al., 2008), stretching as such the VSTM capacity limits to the maximum (Curby & Gauthier, 2010).

In the present study, we provide a further exploration of effects of expertise on VSTM. In particular, we investigate how reading expertise might modify VSTM capacity. To the skilled reader, letters are highly familiar items normally presented in horizontally aligned strings so as to form word units in alphabetic languages such as English and French. Years of reading experience in such languages contribute to the development and utilization of a highly proficient tactic for processing horizontally displayed strings of letters (e.g., Grainger, Tydgat, & Isselé, 2010; Tydgat & Grainger, 2009). Here we examine whether such expertise gained in processing letter strings might influence the storage capacity of unfamiliar random consonant strings in VSTM.

It is thought that the processing of random strings of letters also involves specialized processes (e.g., Grainger & van Heuven, 2003; Tydgat & Grainger, 2009). This is supported by evidence suggesting that the perceptual processes underlying letter string processing are different to those involved in the visual processing of nonletter stimuli. For example, letter perception fails to conform to the acuity gradient principle, according to which there is a processing advantage for visual stimuli presented at fixation that attenuates as the distance from fixation (i.e., eccentricity) increases. In contrast, centrally fixated letter strings show a processing advantage for letters in the central position that drops as a function of eccentricity only to be recovered for letters in the exterior positions (e.g., Hammond & Green, 1982; Ktori & Pitchford, 2008; Mason, 1982; Tydgat & Grainger, 2009). As such, letters in the initial and final positions are processed as efficiently as those in the central position of the letter string. These findings support the argument that letter string processing is influenced by the spatial layout of letters in strings (Grainger & van Heuven, 2003; Tydgat & Grainger, 2009). Indeed, several theoretical and computational frameworks have been put forward attempting to explicate the specialized process of letter string processing (see Davis, 2010; Grainger, 2008; Whitney 2001, for different proposals).

Here we investigated whether the spatial layout of letters making up a string affects the accuracy with which a group of proficient adult readers performed a change-detection task. Change detection is a well-established paradigm in the study of
VSTM (Luck, 2008). In the change-detection task, an array of items (sample array) is presented briefly, and following a retention interval a second item array is displayed (test array). The test array is either identical to the sample array or differs from it in terms of a single item. Participants are required to determine whether or not there is a change between the sample and test arrays while mean response accuracy is recorded. The duration of the retention interval (i.e., 900 ms) is believed to be sufficiently long to exclude contributions from iconic memory (Loftus, Duncan, & Gehrig, 1992; Vogel, Woodman, & Luck, 2001). Finally, combined with an articulatory suppression task, the verbal recoding of visual stimuli is discouraged, and any potential contributions from verbal working memory are thus ruled out (Baddeley, 1986; Besner, Davies, & Daniels, 1981). The change-detection task is currently the dominant paradigm in investigations of VSTM (Luck, 2008).

The starting point of the present research is a finding reported by Predovan et al. (2009) and Prime, Dell’Acqua, Arguin, Gosselin, and Jolicœur, (2010). These authors used an electro-physiological measure of maintenance of information in VSTM, the sustained posterior contralateral negativity (SPCN), to investigate effects of spatial layout on memory for words and nonwords. Disrupting the spatial layout of letter strings was found to modulate the SPCN and affected accuracy on a letter identification task. This finding is to be contrasted with the results of McCollough, Machizawa, and Vogel (2007) showing no influence of spatial layout on the SPCN generated for arrays of simple coloured stimuli. Of course, this comparison is contaminated by a confound with stimulus complexity. The present study provides the first direct comparison of effects of spatial layout on performance to letters and nonletter stimuli of comparable visual complexity in a standard task used to measure VSTM, the change-detection task. Use of this task will allow us to estimate differences in VSTM capacity as a function of stimulus type and spatial layout.

In the present study, participants were presented with visual stimuli that varied in terms of character type (letters vs. symbols), size of stimulus array (3, 5, and 7), and type of display (horizontal, vertical, and circular), in a standard change-detection paradigm (see Figure 1). We predicted enhanced change-detection accuracy for letter arrays presented in horizontal displays as indicative of a specialized mechanism employed during letter string processing.

**Method**

**Participants**
A total of 15 participants drawn from the undergraduate and postgraduate population of the Aix-Marseille University, Marseille, France took part in the study. They all reported normal or corrected-to-normal vision.
**Design**
A repeated measures design was employed, in which the three independent variables were character type (letters and symbols), number of items (3, 5, and 7), and display (horizontal, vertical, and circular). Mean response accuracy in the change-detection task was measured.

**Stimuli**
Stimuli comprised sequences of 3-, 5-, and 7-item arrays. Arrays were constructed using two sets of characters: letters presented in upper case (B, D, F, G, K, L, N, S, and T) and symbols (% , /, ?, @, }, , £, §, and µ). First, a set of sample stimuli was created, followed by a set of test stimuli. The sample stimuli consisted of a random sequence of either letters only or symbols only (without character repetition). As only consonant letters were used, letter arrays were unpronounceable, and care was taken that they did not form acronyms. There were 40 stimuli for each of the 3-, 5-, and 7-item arrays, resulting in a total of 240 stimuli, half of which comprised letter sequences and the other half symbol sequences. The test stimuli consisted of a set of stimulus arrays corresponding to the sample stimuli. Half of the test stimuli were identical to the sample stimuli. The other half contained a character change between the sample and test stimuli by pressing one of two previously assigned keys on the keyboard. Once a response was made, participants were instructed to stop articulating the digits. The next trial began after an intertrial blank interval of 500 ms. All stimuli across the three display types (horizontal, vertical, and circular) were randomly presented within a single block of 720 trials. The option of a short pause was provided following the presentation of 40 trials.

**Results**
A $2 \times 3 \times 3$ within-groups analysis of variance (ANOVA) was conducted on the mean accuracy scores of the participants during the change-detection task (averaged over change and no-change trials combined), with character type (letters or symbols), number of items (3, 5, and 7), and display (horizontal, vertical, and circular) as within-groups variables. Table 1 reports participants’ mean change-detection performance in each of the 18 experimental conditions tested in the present study.

There was a main effect of character type, $F(1, 14) = 28.39, p < .001, \eta^2 = .06$, as overall participants responded to letters more accurately than symbols. There was also a main effect of number of items, $F(2, 28) = 305.36, p < .001, \eta^2 = .60$, revealing the standard capacity limits of VSTM. Furthermore, the main effect of display was significant, $F(2, 28) = 25.02, p < .001, \eta^2 = .06$, as
stimuli presented in horizontal displays were responded to more correctly than those presented in either vertical or circular displays. Most importantly, the three-way interaction between character type, number of items, and display was also significant, $F(4, 56) = 3.091, p < .05, \eta^2 = .01$. The three-way interaction is illustrated in Figure 2.

Follow-up analyses of the three-way interaction revealed that the interaction between number of items and display was significant for letter arrays, $F(4, 56) = 3.99, p < .01, \eta^2 = .03$, but not for symbol arrays, $F(4, 56) = 1.81, p = .14$. As can be seen in Figure 2, the two-way interaction between number of items and display for letter stimuli is mainly driven by variations in the size of the advantage for horizontal displays as a function of the number of items in the display.

Further exploration of the triple interaction shown in Figure 2 revealed that the two-way interaction between character type and number of items was significant for horizontal arrays, $F(2, 28) = 5.65, p < .01, \eta^2 = .04$, but not for vertical displays, $F(2, 28) = 3.15, p = .1$, or circular displays, $F(2, 28) = 0.07, p = .93$. Participants responded more accurately to letters than to symbols with both 5-item and 7-item horizontal displays, $t(14) = 5.45, p < .001$; $t(14) = 3.29, p < .01$, respectively, but not with 3-item horizontal displays, $t(14) = 1.84, p = .089$.

Table 1. Proportion of correct responses to letter and symbol stimuli as a function of the number of items in the array and the different types of display

<table>
<thead>
<tr>
<th>No. of items</th>
<th>Letters</th>
<th></th>
<th></th>
<th>Symbols</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Horizontal</td>
<td>Vertical</td>
<td>Circular</td>
<td>Horizontal</td>
<td>Vertical</td>
<td>Circular</td>
</tr>
<tr>
<td>3</td>
<td>.97</td>
<td>.95</td>
<td>.91</td>
<td>.92</td>
<td>.90</td>
<td>.83</td>
</tr>
<tr>
<td>5</td>
<td>.89</td>
<td>.76</td>
<td>.77</td>
<td>.73</td>
<td>.68</td>
<td>.70</td>
</tr>
<tr>
<td>7</td>
<td>.74</td>
<td>.60</td>
<td>.67</td>
<td>.68</td>
<td>.61</td>
<td>.60</td>
</tr>
</tbody>
</table>

Figure 2. Mean accuracy scores of letters and symbols ($\pm$ 1 standard error of the mean) as a function of the number of items in the array across horizontal, vertical, and circular displays.
Discussion

The present study sought to establish whether expertise in letter processing influences capacity limitations in VSTM. Performance in a change-detection task was examined using item arrays that varied in terms of character type (letters or symbols), number of items (3, 5, and 7), and type of display (horizontal, vertical, and circular).

Results showed an effect of item load on VSTM reflected by the strong linear relation between number of items in the array and reduced response accuracy across both character and display type. In particular, while accuracy was near perfect for 3-item arrays, it declined substantially for 5- and 7-item arrays. This is consistent with previous research reporting a systematic deterioration in change-detection accuracy as the set size of items increases (Luck & Vogel, 1997; Vogel et al., 2001). Furthermore, this finding is a testament to the highly restricted capacity of VSTM confirming its standard storage limits that are estimated to be around four simple items (Cowan, 2001; Phillips, 1974).

Moreover, results revealed an effect of stimulus familiarity, as overall, letters elicited significantly more accurate change-detection responses than symbols. This finding is consistent with previous research also showing a processing advantage for letters over symbols using different experimental paradigms such as visual search and two-alternative forced-choice tasks (e.g., Hammond & Green, 1982; Ktori & Pitchford, 2008; Mason, 1982; Tydgat & Grainger, 2009). Within the scope of the present study, this finding provides evidence that highly familiar items, such as letters, are more accurately encoded in VSTM than unfamiliar items, such as symbols, consistent with the findings of Buttle and Raymond’s (2003) study. Furthermore, the effect of familiarity on VSTM representations supports the argument that VSTM is more accurate for items linked directly to strong and distinct representations in visual long-term memory (VLTM), where visually processed information can be held indefinitely (Luck, 2008).

Importantly, however, the present results provide further evidence that expertise is an essential factor influencing the accuracy with which representations are stored in VSTM. This was reflected by the selective advantage shown for letter over symbol stimuli when presented in horizontal compared to vertical or circular displays. Years of reading experience leads to expertise in processing horizontally aligned strings of letters (for languages that use horizontally aligned scripts), and it is this specific expertise that would be the basis of the selective advantage for horizontally arranged letters strings found in the present study. The fact that the difference in performance to letters and symbols varied significantly as a function of display format and the number of items in the display helps rule out alternative explanations of the present results in terms of differences in familiarity between letters and symbols, or other possible differences such as visual complexity. Since the letter strings in the present study were random combinations of consonants, it is our general experience in processing the specific combinations of horizontally arranged letters that form words that must be driving the pattern of effects that was observed. While horizontal strings of 5 symbols show the standard limitations associated with VSTM (the drop in accuracy relative to 3-symbol strings is a substantial 19%), horizontal strings of 5 letters are relatively spared (the drop in accuracy relative to 3-letter strings is only 8%).

This finding is consistent with previous studies demonstrating effects of expertise on VSTM in the domains of faces (Curby & Gauthier, 2007; Curby et al., 2009; Scolari et al., 2008), cars (Curby et al., 2009), and chess (Chase & Simon, 1973; Freyhof et al., 1992; Gobet & Simon, 1998). However, the underlying mechanism that gives rise to the VSTM advantage of horizontally aligned letter strings might differ from the one attributed to the VSTM advantage elicited by expertise with faces, cars, or chess. Thus, the enhanced VSTM memory capacity observed in chess experts is believed to arise from a process that enables the chunking of information into meaningful long-term memory units, which can be accessed through pointers stored in VSTM (Chase & Simon, 1973; Freyhof et al., 1992; Gobet & Simon, 1998). Similarly, enhanced
VSTM capacity for upright faces and cars, in the absence of an advantage for inverted versions of the same stimuli, is thought to be due to global (holistic) processing of the upright images that is enhanced by expertise with such stimuli (Curby & Gauthier, 2007; Curby et al., 2009; Scolari et al., 2008). On the other hand, chunking mechanisms or global (holistic) processing are unlikely to be the source of the advantage found for random consonant strings in the present work. This is because these random strings of consonants did not form familiar wholes, nor did they contain familiar subparts.

The observed influence of the spatial organization of letter strings on performance in the change-detection task is in line with recent electrophysiological evidence. Specifically, Prime et al. (2010) examined the amplitude of the sustained posterior contralateral negativity (SPCN), which is believed to reflect neural activity specifically related to the maintenance of information stored in VSTM. They showed differences in the amplitude of the SPCN between horizontal and spatially scrambled arrays of letters, thought to reflect variation in the VSTM load induced by these conditions. In particular, it was shown that horizontally aligned letter strings produced a reduced SPCN compared to letters presented in a scrambled fashion, indicating improved maintenance of information in VSTM in the former condition. On the other hand, McCollough et al. (2007) failed to find an influence of spatial layout on the SPCN generated for arrays of simple coloured stimuli.

Together these findings are indicative of a highly efficient mechanism of letter string processing built up over years of reading experience. Indeed, recent theoretical and computational frameworks postulate the utilization of such a specialized mechanism during letter string encoding (e.g., Davis, 2010; Grainger & van Heuven, 2003; Whitney 2001; see Grainger, 2008, for review). For example, Grainger and van Heuven (2003) proposed that letter string processing begins with the operation of a parallel process across a bank of location-specific alphabetic character detectors (referred to as the alphabetic array). They suggested that these alphabetic character detectors perform parallel independent character processing along the horizontal meridian. It is possible that it is expertise in such parallel processing that facilitates the transfer of letter identity information to VSTM, therefore forming the basis of the greater VSTM capacity found for such stimuli in the present study. It is also possible that it is sustained activity in location-specific letter detectors that is the primary basis of the advantage for horizontally arranged letters in the change-detection task. Other types of stimuli, and letters presented in vertical or circular arrays, would require general-purpose mechanisms for maintaining information about object identity and location in VSTM.

Related to this last point, it is possible that the selective advantage for horizontally arranged letter strings in the change-detection task has nothing to do with VSTM, but is simply a reflection of the improved encoding of stimuli in this condition given the limited exposure duration (100 ms) used in the present study. The results of a study using very similar stimuli and with identical exposure durations in a postcued character-in-string identification task would suggest not. In the study in question (Tydgat & Grainger, 2009), participants saw horizontally arranged strings of five characters that could be consonants, digits, or symbols, and they had to identify one out of the five characters at a position cued immediately after stimulus offset (accompanied by a backward mask). Average two-alternative forced-choice accuracy in Experiments 1 and 2 of that study was very slightly higher for letters than for symbols (2–3%), but much smaller than the 16% advantage seen for horizontally arranged five-letter strings over five-symbol strings in the present work.

Finally, some recent research has compared processing of letter strings and symbol strings in dyslexic readers using Tydgat and Grainger’s (2009) postcued character-in-string identification task (Ziegler, Pech-Georgel, Dufau, & Grainger, 2010). This study found that although the dyslexic group showed an overall deficit in letter identification accuracy and no deficit in symbol identification accuracy relative to age-matched controls, the serial position functions were the same as those in normal readers (roughly W-shaped for
letters and inverted U-shaped for symbols). These results suggest that the dyslexic children tested in that study were not impaired at the level of early visuo-orthographic processing, but were more likely impaired at the level of symbol–sound mappings. This would therefore suggest that dyslexic readers should show the same interaction between stimulus type and display type in a change-detection task as that found in the present study—a prediction that clearly merits testing in future research.

Original manuscript received 16 February 2011
Accepted revision received 26 May 2011
First published online 3 October 2011

REFERENCES


472 THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2012, 65 (3)


