Visual working memory for letters varies with familiarity but not complexity

William X. Q. Ngiam[1], Kimberley L. C. Khaw[1], Alex O. Holcombe[1], Patrick T. Goodbourn[1,2]

[1] School of Psychology, The University of Sydney, New South Wales, Australia

[2] Melbourne School of Psychological Sciences, The University of Melbourne, Victoria, Australia

Corresponding Author: William Ngiam william.ngiam@sydney.edu.au

Running Title: Visual working memory varies with familiarity

Note: This preprint was submitted on the 14th March 2018. This manuscript is currently under review for publication and thus is not yet peer reviewed. Please do not copy or cite without author's permission.

Abstract

Visual working memory (VWM) is limited in both the *capacity* of information it can retain and the *rate* at which it encodes that information. We examined the influence of stimulus complexity on these two limitations of VWM. Observers performed a change-detection task with English letters presented in different fonts, or letters from different unfamiliar alphabets. Average *perimetric complexity* (κ)—an objective correlate of the number of features comprising each letter—differed between alphabets. Varying the time between the memory array and mask, we used change-detection performance to estimate the number of items held in VWM (K) as a function of encoding time. For all alphabets, K increased over 270ms (indicating the rate of encoding) before reaching an asymptote (indicating capacity). We found that rate and capacity for each alphabet were unrelated to complexity: Performance was best modelled by assuming that both were limited in terms of number of *items* (K), rather than number of *features* ($K \times \kappa$). We also found a higher encoding rate and capacity for familiar alphabets (~45 items sec^{-1}; ~4 items) than for unfamiliar alphabets (~12 items sec^{-1}; ~1.5 items). We then compared change-detection performance between English letters and an unfamiliar artificial character set matched in complexity. Again, performance was better for the familiar English letters than the unfamiliar characters. We conclude that rate and capacity are determined primarily by familiarity with memoranda. This suggests that binding of letter features is sufficiently rapid to allow encoding of alphabetic stimuli as feature-integrated objects in VWM.

Word Count: 248

Keywords: visual working memory, memory capacity, encoding rate, stimulus complexity, stimulus familiarity.

Introduction

Encoding, storing and manipulating information in *visual working memory* (VWM) is critical for perception and cognition. For example, we use VWM to integrate information across saccadic eye movements (Irwin & Andrews, 1996); to retain information about objects during search and tracking (Luck & Vogel, 2013); and to guide the deployment of attention (Awh & Jonides, 2001). It is also postulated to play a role in higher cognitive functions, as suggested by strong correlations between performance on working-memory tasks and measures of cognitive ability (Cowan et al., 2005; Fukuda, Vogel, Mayr, & Awh, 2010).

Models of VWM architecture

VWM is surprisingly limited in capacity: Typically, only 3 or 4 items can be stored at once. Conflicting findings regarding the influence of stimulus complexity on this capacity limit have inspired different models of VWM architecture. Some have proposed a limit defined strictly by the number of *items* to be memorised. For example, Luck and Vogel reported that VWM performance was invariant when memoranda were conjunctions of various features, such as colour, orientation, size and the presence or absence of a gap (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001). As the requirement to encode additional features did not influence change-detection performance, they concluded that each object is stored in a VWM ‘slot’ with its features already integrated. In contrast, Alvarez and Cavanagh (2004) reported different VWM capacities for stimulus sets differing in complexity. They defined the complexity of a set as its visual-search rate, the additional time taken to find a target with each additional object in a visual-search array. Visual-search rate was strongly (indeed, almost perfectly) correlated with the inverse of VWM capacity for each stimulus set: That is, VWM capacity was lower for more complex objects. Such findings suggest a resource model of VWM whereby storage of more complex objects requires more of some limited resource. Visual-search rates are better predictors of VWM capacity for short stimulus presentations than for longer ones (Eng, Chen, & Jiang, 2005), suggesting that stimulus complexity might impact VWM capacity by limiting the rate of perceptual encoding.

Awh, Barton and Vogel (2007) contend that the apparent differences in VWM capacity observed by Alvarez and Cavanagh (2004) were due not to stimulus complexity *per se*, but rather to an increase in confusion errors when memory and test objects were more visually similar (high *sample–test similarity*). They compared change-detection accuracy when the change was ‘within-category’, such as from a shaded cube to another, to when the change was ‘cross-category’, such as from a shaded cube to a Chinese character. Change-detection accuracy with within-category changes decreased as complexity increased, but change-detection accuracy with cross-category changes was equivalent to change-detection accuracy with low complexity items (simple colours). By their account, the effect of complexity on change-detection performance comes when memory and test objects must be compared, rather than during encoding. Note that these findings do not directly contradict Alvarez and Cavanagh’s basic claim that change-detection performance is influenced by the complexity of stimuli. A more complex object may be represented in VWM with fewer intact features, or at a lower resolution. A degraded representation of a Chinese character might be easily distinguishable from a coloured square, but may not be distinguishable from another character with similar features.

In addition to being limited in capacity, VWM also appears to be limited in the rate at which representations can be encoded. Vogel, Woodman and Luck (2006) quantified the rate of encoding objects into VWM by manipulating the delay or *stimulus onset asynchrony* (SOA) between the presentation of the memory array and a backward-mask array in a change-detection task. Vogel et al. (2006) found performance on the change-detection task improved steadily for SOAs up to 200 ms, before reaching an asymptote. Prior to the asymptote, each colour block took approximately 50 ms to encode; beyond the asymptote, capacity was approximately 2.5 blocks.

Considering encoding rate also addresses a potential confound in standard approaches to measuring VWM capacity. Typically, the same temporal relationship between memory, mask and test arrays is maintained throughout an experiment. If the time available to encode objects is too brief, the capacity of VWM might be underestimated. Furthermore, if encoding rate varies with the complexity of the memoranda, sets of more complex objects will require more time to saturate VWM capacity. If this were the case, apparent differences in VWM capacity for objects of different complexity—such as those reported by Alvarez and Cavanagh (2004)—might be an artefact of interrupting encoding before more complex objects have had sufficient time to fill VWM. As Vogel et al. (2006) used a single stimulus set, it is unknown whether the rate of encoding into VWM changes with stimulus complexity. An object with more features may take longer to encode, as suggested by Eng et al. (2005); alternatively, consolidation into VWM may be independent of the object's complexity.

Complexity and familiarity

The definition of *complexity* has been inconsistent in the VWM literature, and differing definitions may have led to conflicting findings. Luck and Vogel (1997) manipulated complexity by conjoining visual features across different dimensions (e.g. colour and orientation). Varying complexity by adding features in *different* dimensions will have no effect if VWM comprises independent stores for different feature dimensions. Alvarez and Cavanagh (2004) estimated object complexity using a visual-search task; yet, a correlation between visual-search performance and VWM capacity should be expected because visual search relies on VWM processes. Emrich, Al-Aidroos, Pratt and Ferber (2010) reported that visual-search rate decreased when a memory load that exhausted VWM capacity was placed on observers, but not when the load was below VWM capacity. They suggest that VWM inhibits returning to previously searched items, but this is limited by the number of items that can be held in VWM capacity (see Le-Hoa Võ & Wolfe, 2015). Visual-search rates thus may not index stimulus complexity *per se* but rather a range of factors that influence both change-detection and visual search performance, such as item similarity and familiarity.

Although objective measures of visual complexity exist, to our knowledge they have not been applied to the study of VWM. For example, *perimetric complexity* is defined as the square of the combined inside and outside perimeters of a letter, divided by its area (Attneave & Arnoult, 1956). Letter identification efficiency stands in a nearly perfect negative linear relationship with perimetric complexity, suggesting that letter identification is mediated by detectors of simple features: As letters increase in complexity, they are identified increasingly inefficiently because more features must be combined together (Pelli, Burns, Farell, & Moore-Page, 2006). An object's perimetric complexity thus provides a useful proxy for the number of features it contains. Importantly, this allows complexity to vary without the addition of extra feature dimensions such as colour.

Familiarity has also been shown to influence consolidation and storage in VWM. For example, higher VWM capacities have been found for famous faces over unfamiliar faces (Jackson & Raymond, 2008) and for Pokémon (cartoon characters) from the original generation over the recent generation among participants who reported familiarity with the original generation (Xie & Zhang, 2016). Additionally, the encoding rate for Pokémon was faster for people familiar with them compared to those who were not (Xie & Zhang, 2017). However, as these studies do not control for or measure stimulus complexity, the degree to which these effects of familiarity are independent of stimulus complexity is unknown.

The present study

In the present study, we examined the influence of stimulus complexity and familiarity on the capacity and encoding rate of VWM. We used a variant of Vogel et al.'s (2006) change-detection paradigm, whereby the number of items held in VWM was measured as a function of the amount of time allowed for stimulus encoding. In Experiment 1, we used letters of the English alphabet, and varied complexity by presenting the letters in four different fonts. To anticipate out results, we did not find an influence of stimulus complexity, but what appeared to be an effect of familiarity. In Experiment 2, we used characters from four alphabets that were equally unfamiliar to our participants, to isolate any possible effect of stimulus complexity independent of familiarity. In Experiment 3, we compared English characters with the Brussels Artificial Character Set (BACS), which is designed to have the same featural information, such as the number of junctions, strokes and terminations as English characters (Vidal, Content, & Chetail, 2017). This allowed us to examine whether familiarity with English letters led to any differences in VWM performance compared to the unfamiliar but similarly complex BACS letters.

We considered four general models of VWM architecture. Both encoding and capacity limits in VWM might best be described in terms of *objects*, or in terms of *features*. If feature integration is a limiting factor in the process of consolidation into VWM, letters from a more complex alphabet (which contain more features) will be encoded at a slower rate. Alternatively, feature integration may not limit VWM consolidation; in this case, encoding rate for items will not vary with stimulus complexity. Likewise, if the number of features that can be stored limits VWM capacity, fewer items will be stored from more complex alphabets. Alternatively, the number of features may not limit VWM capacity; in this case, the number of items that can be stored will not vary with stimulus complexity.

Figure 1 shows predictions from the four ensuing models of VWM encoding and storage. Each panel shows the number of objects stored in VWM (*K*) as a function of encoding time—that is, the stimulus onset asynchrony (SOA) between the memory array and the mask array. The initial gradient of each function is determined by encoding rate, and the asymptote reflects VWM capacity (Vogel et al., 2006). Luck and Vogel (1997) found that VWM capacity did not vary with the complexity of stimuli. This is consistent with the assumption that VWM storage comprises a fixed number of feature-integrated object files, such that its capacity is independent of item complexity. Of our four models, this is consistent with both I_rI_c (*rate in items, capacity in items*; Figure 1a) and F_rI_c (*rate in features, capacity in items*; Figure 1c), which have an identical asymptote for each alphabet. In contrast, Alvarez and Cavanagh (2004) found that VWM capacity varied with item complexity, such that fewer items can be stored with increasing object complexity. Of our four models, this is consistent with both I_rF_c (*rate in items, capacity in features*; Figure 1b) and F_rF_c (*rate in features, capacity in features*; Figure 1d), with a different asymptote for each alphabet that is predicted by item complexity. The F_r models (*rate in features*) predicts varying encoding rates according to the perimetric complexity of the alphabet, whereas I_r models (*rate in items*) predict perimetric complexity has no influence on encoding rates of an alphabet.

General Methods

Apparatus

Stimuli were generated using MATLAB (The MathWorks, Natick, MA) and PsychToolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). They were shown on a Trinitron G520 CRT monitor (Sony Corporation, Tokyo, Japan) on a uniform grey background of luminance 51 cd m^{-2}. The spatial resolution of the monitor was 1024 × 768 pixels, and the refresh rate was 100 Hz. The monitor was gamma-corrected independently for each of the red, green and blue phosphors; we measured gamma functions using a Spyder4ELITE photosensor (Datacolor, Lawrenceville, NJ) and applied corrections using the PsychToolbox imaging pipeline. Participants were seated in a dark room, with a chin rest to maintain a viewing distance of 57 cm from the monitor.

Stimuli

Each letter was drawn in black, within a white circular aperture. An aperture subtended 1.8°, and a letter subtended a maximum of 1.5° on vertical and horizontal axes. Each aperture was positioned on the circumference of an imaginary circle such that each was centred 4.0° from the fixation point. Apertures were equally spaced around the circle, with a random rotation applied to the circle of apertures on each trial.

Procedure

The general procedure for each trial is shown in Figure 2. A warning tone sounded at the beginning of each trial. At the same time, a fixation point appeared with two randomly selected digits (1–9) on either side (centred 3.2° to the left and right). Participants were instructed to repeat these numbers aloud throughout the trial. This articulatory suppression procedure is used to interrupt verbal encoding of the letter stimuli (Besner, Davies, & Daniels, 1981), which might otherwise allow participants to rely on systems other than visual working memory to perform the task.

The numbers disappeared after 1000 ms, leaving a blank screen with the fixation point for 1000 ms, after which the memory array was shown for 100 ms. The succeeding dynamic mask array contained phase-scrambled transformations of all letters in the set, displayed in each circular aperture. Scrambling the Fourier phase spectrum of an image retains the spatial-frequency content while destroying overall form information. The mask was shown for 200 ms, comprising 10 different transformations displayed for 20 ms each. Mask onset occurred 20, 30, 60, 100, 170, 290 or 500 ms after offset of the memory array. The stimulus onset asynchronies (SOAs)—that is, the set of delays between the onset of the memory array and the onset of the mask—were therefore 120, 130, 160, 200, 270, 390 or 600 ms. By concentrating SOAs at the lower end of the range, we could measure encoding rate more precisely.

Onset of the test array always followed 1000 ms after onset of the memory array. In each test array, one letter was randomly changed to one of the remaining letters in the set. Participants were required to identify where the change had occurred by using the computer mouse to click on the circular aperture of the changed letter.

Feedback was provided: A high tone was played after a correct response, or a low tone after an incorrect response. The next trial commenced 1000 ms after feedback.

Each participant completed 16 blocks per experiment—four blocks for each of four alphabets—over two or three weeks. Each block lasted approximately 20 mins, and contained 210 trials in total (30 trials at each of seven SOAs). Within each block, participants were prompted to take a short break after completing each set of 70 trials. In total, each participant completed 3360 trials per experiment: 120 at each of the seven SOAs, for each of the four alphabets.

Estimating perimetric complexity

Perimetric complexity (κ) was defined as the inside and outside perimeter of the stimuli (P_{total}) squared, divided by the area (A), all divided by 4π,

$$\kappa = \frac{{P_{total}}^2}{4\pi A}. \quad \text{(Equation 1)}$$

Alphabets and fonts were selected to span a range of the perimetric complexity values reported by Pelli et al. (2006). Note, however, that their definition of perimetric complexity excludes the division by 4π that we introduce here. The simplest shape, with the lowest ratio of the squared perimeter to area, is a circle (squared perimeter = $4\pi^2r^2$, area = πr^2, ratio = 4π). By dividing by 4π, our measure of complexity assigns the value of 1 to a circle, the simplest shape, which may ease interpretation. Because the perimetric complexity of a stimulus depends on the effective resolution of the display (Watson, 2011), and the resolution of our display likely differs from that of Pelli et al. (2006), we calculated values for this experiment using a MATLAB program based on the algorithm specified by Pelli et al. (2006)[1].

Equipped with estimates of perimetric complexity, we could also assess the number of *features* stored in VWM. Using κ as a proxy for the average number of features (up to a proportionality constant) contained within a letter of an alphabet, the number of stored features is the product of the number of stored items and complexity,

$$\text{Number of features} = K \times \kappa \qquad \text{(Equation 2)}$$

Estimating items stored in VWM

Pashler's (1988) formula was adapted for our change-detection task, as we asked participants to identify the *location* of the change rather than whether or not a change occurred. If the participant encodes, on average, a given number of objects in VWM (*K*) out of the total number of objects in the array (*N*), we make the straightforward assumption that they will detect the location of change on *K* out of *N* trials. On the remaining (*K*−*N)* trials, when they do not detect the location of change, the participants will have to guess from the array. When guessing, participants may randomly select one target from the array, such that they will have a 1 out of *N* probability of selecting the correct letter. Thus, the proportion of correct responses (*P*) will be

$$P = \frac{K}{N} + \frac{1}{N}\left(1 - \frac{K}{N}\right). \qquad \text{(Equation 3)}$$

Rearranged to estimate *K*,

[1] The algorithm specified by Pelli et al. (2006) can be found as a footnote on page 4652 of the article.

$$K = \frac{PN-1}{\left(1-\frac{1}{N}\right)}. \qquad \text{(Equation 4)}$$

We refer to Equation 4 as the *random-guessing* formulation. Of course, this assumes that when a participant cannot identify the location of the change, she selects randomly from the array. If a participant is able to inform her guess by inferring that it did not occur at any of the successfully encoded locations, she will have a 1 out of (N−K) probability of selecting the correct letter, such that

$$P = \frac{K}{N} + \left(\frac{1}{N-K}\right)\left(1-\frac{K}{N}\right). \qquad \text{(Equation 5)}$$

When rearranged to estimate K,

$$K = PN - 1. \qquad \text{(Equation 6)}$$

We refer to Equation 6 as the *informed-guessing* formulation. For simplicity, the analyses presented here have been conducted with the random-guessing formulation (Equation 6). The two formulations produce very similar estimates of K, and the conclusions do not change when the informed guessing formulation is used.

For each alphabet, we considered K as a function of SOA. The slope of the initial part of this function is taken to be the encoding rate (Vogel et al., 2006). The function was expected to reach an asymptote at longer SOAs; this asymptote is interpreted as the capacity of VWM. We estimated encoding rate and capacity by fitting a combination of two linear functions using a least-squares procedure. The first function was restricted to pass through the origin, with gradient as a free parameter (encoding rate). The second was restricted to have zero gradient, with the constant y-value as a free parameter (capacity). The domain of the first function was restricted to x-values below where the two functions intersected; the domain of the second was restricted to *x*-values above that point. The result was a function of the general shape shown in Figure 1a.

Modelling and model comparison

We used likelihood-maximisation procedures to determine which of the four models depicted in Figure 1 best fit each participant's data. Maximum-likelihood estimation fits a model to a set of observed data by finding those parameters that maximise the probability of the data. Each of the models has only two parameters, *rate* (slope of the initial linear function) and *capacity* (asymptote). In some models, one or both parameters were dependent on number of features (e.g. the rate parameter in the F_rI_c model). In these cases, the parameters were multiplied by the complexity value for each alphabet. For each participant, we fit each of the four models separately. We ran a hundred iterations of each fit, starting with different random parameter estimates. To facilitate interpretation and comparison, likelihood values of the best fits for each model were converted to Bayes Information Criterion (BIC) values (Schwarz, 1978). We used these BIC values to select the best-fitting model for each participant's data, with a lower BIC indicating a better fit (Kass & Raftery, 1995). The best-fitting model thus indicates whether the appropriate measure of encoding rate and capacity of VWM is items or features.

Experiment 1

In Experiment 1, encoding rate and capacity of VWM were estimated for letters from four different English fonts. Vogel et al. (2006) found that for coloured squares, encoding rate was 0.02 items per millisecond (one item every 50 milliseconds), and capacity was approximately 2.5 items. Here, we use four fonts differing in average perimetric complexity, which has been shown to predict the efficiency with which letters are identified (Pelli et al., 2006). This allowed us to examine whether a more complex letter takes longer to consolidate into VWM, and whether there is a smaller capacity in VWM for more complex letters relative to simpler ones.

Methods

Participants

Four males and one female completed this experiment: two authors (WN & PG) and three colleagues unaware of the aims of the experiment. All participants reported normal or corrected-to-normal visual acuity.

Stimuli

On each trial, six letters were shown in each of the memory and test arrays. The letters came from four different English font alphabets: Helvetica Bold (lowercase), Courier (lowercase), Bookman Old Style (uppercase) and Künstler Script (uppercase). The letters were generated using TrueType fonts from Apple OSX 10.7.5 and converted to greyscale bitmap images (see Figure 3). We selected these fonts to produce a spread of perimetric complexity; based on Pelli et al.'s (2006) calculations, initial estimates for each font were 5.3, 8.0, 11.1 and 35.9, respectively.[2] Each alphabet comprised 20 different letters. The letters C, F, I, N, V and W were removed, as pilot studies indicated they were the most commonly confused with other letters. For each array, six letters were selected randomly without replacement. We used a set size of six because pilot studies indicated that smaller set sizes produced ceiling effects.

[2] Pelli et al. (2006) report these values as 67, 100, 139 and 451. Here, we normalise these values by dividing by the complexity of the simplest possible object (4π).

Results and discussion

Perimetric complexity

We first calculated the mean perimetric complexity of each alphabet when displayed on our equipment. The means and the standard deviations are shown in Table 1. Absolute values changed somewhat, but the ordering of complexity was maintained from Pelli et al. (2006).

Table 1

Mean Perimetric Complexity of Alphabets in Experiment 1

Alphabet	Estimated Perimetric Complexity*	Calculated Perimetric Complexity	
		Mean	SD
Helvetica	5.3	6.8	1.1
Courier	8.0	7.4	1.0
Bookman	11.1	10.3	1.6
Künstler	35.9	27.1	5.4

* Perimetric complexity value reported in Table A of Pelli et al. (2006), divided by 4π.

Change-detection performance

Figure 4a shows accuracy (percentage correct) pooled across participants as a function of SOA, for each of the four alphabets. Figure 4b shows the number of items (K) as a function of SOA for each alphabet; and Figure 4c shows the estimated number of features ($K\kappa$) as a function of SOA for each alphabet. Performance increased monotonically before reaching a pleateau after about 270ms for all alphabets. Although our experiment was primarily designed for modelling at the level of the individual participant, we conducted preliminary analyses using one-way repeated measures ANOVAs of rate and capacity, with alphabet as a within-subjects factor. There was a significant effect of alphabet on encoding rate ($F_{3,\,12} = 8.2$, $p = .003$). Using Tukey's HSD, we conducted pairwise comparisons between encoding rates for each alphabet. The mean encoding rate (in items per second) for Künstler ($M = 16.5$, $SD = 9.2$) was significantly different ($p < .05$) from the mean encoding rate for Helvetica ($M = 41.9$, $SD = 20.0$), Courier ($M = 42.5$, $SD = 16.3$), and Bookman ($M = 48.5$, $SD = 22.9$). There were no other significant differences between encoding rates. There was a significant effect of alphabet on capacity ($F_{3,\,12} = 199.1$, $p < .001$). The capacity of VWM (in items) for Künstler ($M = 1.1$, $SD = 0.7$) was significantly different ($p < .05$) from the capacity of VWM for Helvetica ($M = 4.2$, $SD = 0.6$), Courier ($M = 3.9$, $SD = 0.7$), and Bookman ($M = 4.16$, $SD = 0.7$). There were no other significant differences between capacities. The low capacity estimate for Künstler (1.1 items) was not due to floor effects, as performance was significantly above chance for each individual across all SOAs (all observed $\chi^2 > 9.33$, all $p < 0.002$).

Modelling

The maximum-likelihood procedure was conducted to fit each of the candidate models (Figure 1) separately for each participant. Figure 5a shows the ranking of BIC values of each model, for each participant. A lower value indicates a better fit of the model to the data. For all participants, the best-fitting model was I_rF_c. According to this model, encoding rate is set in terms of items, and VWM capacity is set in terms of features. However, visual inspection of the fits showed that the extremely different values we obtained for Künstler precluded any model from fitting well. We thus repeated the procedure with Künstler excluded (Figure 5b).

BIC values were considerably lower for all models when the Künstler font was excluded. A difference of 10 between BIC values is typically interpreted to indicate ‘decisive’ evidence in favour of the model with lower BIC (Kass & Raftery, 1995). BIC values reduced by more than 1000 when Künstler was excluded, indicating overwhelmingly better fits to the data. With Künstler excluded, the best fitting model for each participants’ observed data was the I_rI_c model, which sets both encoding rate and capacity according to number of items. For every participant, BIC values for the I_rI_c model were at least 10 lower than for any other model. This is a strong indication that for these fonts, I_rI_c most accurately represents the workings of VWM. That is, for the Helvetica, Courier and Bookman fonts, the best model is one that quantifies encoding rate and VWM capacity in terms of items, regardless of their complexity.

Discusssion

Experiment 1 was conducted to determine the influence of perimetric complexity on the encoding rate and capacity of VWM. The shape of the encoding function observed by Vogel et al. (2006) for colour blocks was replicated with letter stimuli from different fonts. For all alphabets, performance on the change-detection task was poor when the mask was presented almost immediately following the memory array. As the mask was delayed further, increasing the amount of time for VWM encoding, performance gradually improved before reaching an asymptote. Encoding rate and VWM capacity did not differ significantly between Helvetica, Courier and Bookman, but was significantly different for Künstler. Among the three similar fonts, encoding rate did not decrease with increasing perimetric complexity, as would be predicted if feature binding was a limiting factor in VWM encoding. Similarly, the capacity of VWM did not decrease with increasing perimetric complexity, as would be predicted by resource models of VWM. Estimates of capacity of VWM for these fonts were approximately equal to the capacity estimate reported by Alvarez and Cavanagh (2004) of 3.7 letters.

Differences in the encoding rate and the VWM capacity for Künstler may be due to its greater complexity, but given the lack of variation among the other fonts, it may instead have been caused by participants' lack of familiarity. Unlike the other fonts used in Experiment 1, Künstler is an uncommonly used decorative script, which is likely to be unfamiliar to observers. In their study of the relationship between perimetric complexity and identification efficiency, Pelli et al. (2006) discarded the first 400 trials using Künstler script, citing participants' initial unfamiliarity with the letters. Maximum-likelihood estimation including Künstler indicated that I_rF_c, which defines encoding rate in terms of items and VWM capacity in terms of features, was the best-fitting model. However, the procedure appeared to be biased by the considerably different pattern of performance observed for Künstler. Excluding Künstler reduced BIC values by more than 1000, indicating a much better fit than any model with Künstler included. Here, maximum-likelihood estimation indicated that the best-fitting model was I_rI_c, which quantifies encoding rate and capacity of VWM in terms of items. This suggests that the perimetric complexity of the stimuli does not influence the encoding rate or the capacity of VWM for these stimuli.

Alternatively, it is possible that using familiar stimuli precluded any effect of perimetric complexity on VWM processes. As Helvetica, Courier and Bookman are commonly used fonts, the visual system may more efficiently encode and store these letters in VWM. Familiarity may also allow participants to exploit diagnostic features that distinguish letters without requiring all features to be encoded. Unfamiliar stimuli may be required to reveal an effect of perimetric complexity on the encoding rate and capacity of VWM.

Experiment 2

In Experiment 1, we found little to no effect of perimetric complexity on VWM encoding rate or capacity. One limitation of that experiment was that all fonts except Künstler were familiar to participants. It is unclear whether the latter yielded different results because of its higher perimetric complexity, or rather because it was unfamiliar. Here, we sought to examine the effect of perimetric complexity with unfamiliar sets of letters from languages other than English. In Experiment 1, complexity was partially confounded with familiarity; here, we can isolate any effect of complexity independent of familiarity with the memoranda.

Methods

Participants

Five participants (two female) completed the experiment. They included two of the authors (WN & PG) and three participants who were unaware of the aims of the study and had not completed Experiment 1. All participants reported normal or corrected-to-normal visual acuity, as well as unfamiliarity with the alphabets used in the experiment.

Stimuli

Stimuli were presented in the same manner as for Experiment 1. However, only four symbols were presented in the memory and test arrays (Figure 6a). Presenting six symbols, as in Experiment 1, produced floor effects in pilot data. Four character sets were drawn from Braille, Hebrew, Arabic and simplified Chinese (see Figure 6b). For the Braille alphabet, we generated the symbols using the *Apple Symbols* font; for the Hebrew alphabet, the *Times New Roman* font; for the Arabic alphabet, the *Al Bayan* font; and for the Chinese character set, the *Yung* font.[3] We selected these fonts to span a wide range of perimetric complexity estimates, as reported by Pelli et al. (2006). For each set, we selected 20 characters at random from the 22 (Hebrew), 26 (Braille and Chinese) or 28 (Arabic) available in the chosen fonts. The items presented in each array were selected randomly without replacement from the 20 characters comprising each set.

Results and discussion

Perimetric complexity

The means and standard deviations of perimetric complexity for each character set are shown in Table 2. On our equipment, the order of perimetric complexity values differed from those calculated by Pelli et al. (2006), but the mean for the most complex alphabet (Chinese, 11.9 features) was still almost double that for the simplest (Braille, 7.6 features). Most notably, the perimetric complexity for Braille was considerably higher than previously reported. Pelli et al.'s Braille symbols comprised abutting square elements, whereas our symbols comprised separated circular elements. Pelli et al.'s letters have lower complexity primarily because all adjacent square elements are merged, reducing the total perimeter, and complexity is proportional to the square of total perimeter.

[3] Downloaded from http://psych.nyu.edu/pelli/software.html.

Table 2

Mean Perimetric Complexity of Character Sets in Experiment 2

Character Set	Estimated Perimetric Complexity	Calculated Perimetric Complexity	
		Mean	SD
Braille	2.2	7.6	1.4
Hebrew	7.2	6.6	1.6
Arabic	10.9	8.8	2.3
Chinese	15.8	11.9	2.6

* Perimetric complexity value reported in Table A of Pelli et al. (2006), divided by 4π.

Change-detection performance

Figure 7a shows the mean accuracy (percentage correct) as a function of SOA for each alphabet. Figure 7b shows K as a function of SOA for each alphabet; and Figure 7c shows the estimated number of features ($K\kappa$) as a function of SOA for each alphabet. As in Experiment 1, a combination of two lines was fitted to estimate the encoding rate and capacity for each alphabet, separately for each participant.

As in Experiment 1, we conducted a preliminary one-way repeated-measures ANOVA, which revealed a significant effect of character set on encoding rate ($F_{3,\,12}$ = 13.2, p < .001). The encoding rate (in items per second) for Braille (M = 7.1, SD = 5.0) was significantly different (p < .05) from the mean encoding rate for Arabic (M = 15.1, SD = 8.4) and for Hebrew (M = 15.0, SD = 8.6). We found no other significant differences. The mean encoding rate for Chinese was 11.3 (SD = 6.9). There was no significant main effect of alphabet on VWM capacity ($F_{3,\,12}$ = .7, p = .59). Mean capacity (in items) for Braille was 1.5 (SD = 0.6); for Hebrew, 1.8 (SD = 0.8); for Arabic, 1.7 (SD = 0.5); and for Chinese, 1.6 (SD = 0.6).

Modelling

Figure 8 shows the BIC values of each model for each participant. For four out of the five participants, the best-fitting model was I_rI_c, which quantifies both encoding rate and VWM capacity in terms of number of items. For one participant, there is ‘decisive’ evidence in favour of the model compared to the next most likely model (ΔBIC > 10); for two more, there is ‘strong’ evidence (6 < ΔBIC < 10); and for one, ‘positive’ evidence (0 < ΔBIC < 6; Kass and Raftery, 1995). For the remaining participant, BIC values differ by only 3 between the best and worst model. Overall, we interpret this as good evidence in favour of the I_rI_c model.

Discussion

Experiment 2 was conducted to determine whether familiarity with the stimuli masked an influence of perimetric complexity on VWM limits in Experiment 1. To do this, we replicated Experiment 1 with letters from alphabets that were unfamiliar to the participants.

Mean encoding rate and VWM capacity for each of the character sets in Experiments 1 and 2 is shown in Figure 9. Most notably, encoding rate and capacity for Künstler in Experiment 1 were highly similar to those of the unfamiliar alphabets in Experiment 2. Encoding rate for Künstler was 16.5 items per second, which lies in the upper end of the range for unfamiliar alphabets. VWM capacity for Künstler (1.1 items) is similar to VWM capacity estimates for unfamiliar alphabets, which varied between 1.5 and 1.8 items. The similarity of these values suggests that differences in encoding rate and VWM capacity between Künstler and the more common English fonts in Experiment 1 were likely due to familiarity with the stimuli. Pairwise comparisons indicated the encoding rate of Braille symbols was slower than the encoding rate of Hebrew and Arabic characters; however, the magnitude of the difference is small compared to the difference between familiar and unfamiliar character sets.

Maximum-likelihood estimation indicated that I_rI_c was the best-fitting of our candidate models (Figure 8). The model quantifies both encoding rate and the capacity of VWM in terms of items, rather than in terms of features. As in Experiment 1, these results suggest that encoding rate and VWM capacity for unfamiliar stimuli do not vary proportionally with the number of features to be encoded (as indexed by perimetric complexity).

Experiment 3

Considering the results of Experiments 1 and 2 together, familiarity with memoranda appears to have a pronounced effect on VWM limits. In Experiment 3, we examined the effect of familiarity directly by matching the complexity of our stimuli. To do this, we compared VWM limits for English letters to VWM limits for BACS characters (Vidal et al., 2017), which were designed to match the features of English letters but are novel to our participants. The BACS characters contain the same number of strokes, junctions and terminations as English letters; in addition, we selected an English font that best matched the perimetric complexity of the BACS character set. This allows us to isolate the effect of familiarity while controlling for stimulus complexity.

Methods

Participants

Ten participants (6 males) with normal or corrected-to-normal vision completed the experiment. All were naïve to the aims, and none had participated in Experiments 1 or 2.

Stimuli

Stimuli were presented in the same manner as Experiment 2. English letters were Courier New lowercase, generated using TrueType fonts from Apple OSX 10.7.5. Artificial letters were the serif BACS-2 character set[4]. The BACS equivalents of the most commonly confused English letters (C, F, I, N, V and W) were excluded, leaving 20 matched characters. The items presented in each array were selected randomly without replacement from the set of 20 for each alphabet (Figure 10).

Results

Perimetric complexity

The perimetric complexity of the set of Courier New letters (M = 11.6, SD = 1.7) was not significantly different from the set of BACS letters (M = 10.8, SD = 2.4), $t(38) = 1.21$, $p = .233$. Thus the BACS and Courier New sets are comparable in perimetric complexity, and are matched on the number of junctions, strokes and terminations in each letter.

Change-detection performance

Figure 11a shows the mean accuracy (percentage correct) as a function of the SOA for each alphabet. Figure 11b shows K as a function of SOA for each alphabet; and Figure 11c shows the estimated number of features ($K\kappa$) as a function of SOA for each alphabet. As in Experiments 1 and 2, a combination of two lines was fitted to

[4] Downloaded from https://osf.io/s4738/.

estimate the encoding rate and capacity for each alphabet, separately for each participant. Mean encoding rate (in items per second) was significantly faster for Courier New letters (M = 22.8 SD = 6.1) than for BACS letters (M = 15.7, SD = 6.5), $t(9) = 3.76$, $p < .01$. Additionally, the capacity was significantly higher for Courier New letters (M = 2.5, SD = 0.5) than for BACS letters (M = 1.9, SD = 0.3), $t(9) = 5.17$, $p < .01$.

Discussion

Experiment 3 was conducted to examine the effect of familiarity on VWM limits by matching the complexity of different letter sets. To do this, we used the BACS, an artificial character set designed to match the features of the English letters, such as the number of junctions, strokes and terminations. The BACS characters were not significantly different from the Courier New characters in perimetric complexity, an objective estimate of the number of features in an item. We found that encoding rate and capacity for the familiar English character set was significantly higher than the unfamiliar BACS character set. This is consistent with the results of Experiments 1 and 2, which showed familiarity to be the primary determinant of VWM encoding rate and capacity.

General Discussion

The current study examined the influence of visual complexity and familiarity on the encoding rate and capacity of VWM for alphabetic stimuli. We used the perimetric complexity of letters as an objective, intrinsic measure of stimulus complexity, which estimates the number of basic visual features it contains. Our results suggest that encoding rate and capacity of VWM are not influenced by stimulus complexity. The best-fitting models of VWM architecture in Experiments 1 and 2 quantified the encoding rate and capacity of VWM in terms of items, rather than features. However, marked differences in encoding rate and the capacity are evident between familiar and unfamiliar fonts or alphabets. The commonly used English fonts of Helvetica, Courier and Bookman were encoded significantly faster than letters from the decorative Künstler font, and from the unfamiliar Braille, Hebrew, Arabic and Chinese alphabets. Additionally, the capacity of VWM for the common fonts was significantly higher than for the unfamiliar fonts (see Figure 9). This finding was replicated with familiar English letters rendered in Courier New font and unfamiliar letters from the Brussels Artificial Character Set. Although the two sets of letters were equivalent in visual complexity, the familiar English font showed a significantly faster encoding rate and higher VWM capacity.

Complexity and VWM capacity

Previous findings regarding the influence of stimulus complexity on the capacity of VWM have been mixed. Luck and Vogel (1997) reported that the capacity of VWM did not decrease for objects that were conjunctions of a greater number of features; on the other hand, Alvarez and Cavanagh (2004) reported varying capacity estimates, proportional to the visual information load of the item. In the present study, we defined stimulus complexity as perimetric complexity, an objective estimate of the number of features contained within a stimulus. We found that the capacity of VWM was invariant with the perimetric complexity of the memoranda. This suggests that additional features in the same dimension do not reduce VWM capacity, as Luck and Vogel (1997) found for additional features in different dimensions. Alvarez and Cavanagh's findings differ from ours, suggesting that their visual-search measure of visual information load does not index low-level visual features. We note that perimetric complexity is only weakly correlated with subjective ratings of complexity

(Pelli et al., 2006). Subjective metrics (which may be influenced by familiarity) may be more predictive of VWM limits than the objective measure used in the current study.

Complexity and VWM encoding rate

In addition to VWM capacity, the current study examined the influence of stimulus complexity on the encoding rate of VWM. The best-fitting of our candidate models for change-detection performance was I_rI_c, which quantifies encoding rate in terms of objects, rather than features. The observation that encoding rate is invariant with the complexity of the stimulus suggests that the number of features to be encoded does not temporally limit the encoding process. Rather, the rate of VWM encoding appears to reflect the consolidation of feature-integrated objects.

Encoding objects into VWM has been modelled as a two-stage process of individuation and consolidation (Wutz & Melcher, 2013). The individuation of an object—the process of binding features into a discrete unit—precedes the consolidation of that object into VWM. The efficiency with which letters are identified is inversely proportional to their perimetric complexity, such that letters seem to be identified via their constituent features (Pelli et al., 2006). It is thus reasonable to expect that the integration of features will be limited by perimetric complexity; however, this may occur sufficiently rapidly such that it does not place a bottleneck on VWM encoding. Wutz and Melcher (2013) found that enumeration performance for two-line drawings, a measure of object individuation, was limited when objects were presented for approximately 70 ms, but not when presented for approximately 120 ms. A binding limit of 120 ms has been reported for four-letter English words, along with a limit of 140 ms for pseudowords and 170 ms for pairs of

Chinese characters (Holcombe & Judson, 2007), suggesting that the binding of features for multiple letters may occur over this timescale. As the lowest SOA we used in the current study was approximately 120 ms, the features of each object already may have been bound into a unitary representation. Encoding rate would thus reflect the consolidation of feature-integrated units.

Familiarity

The encoding rate and capacity of VWM for familiar alphabets was higher than for unfamiliar alphabets. In Experiment 1, encoding rate for English letters of common fonts was approximately 45 items per second (22 ms per item), and capacity was approximately four items. This is consistent with the encoding rate of 20–30 ms per item previously reported by Gegenfurtner and Sperling (1993) in a cued recall task for English letters, and numerous findings that capacity is approximately four items (Awh et al., 2007; Cowan et al., 2005; Luck & Vogel, 1997). In Experiment 3, encoding rate for the Courier New font was approximately 23 items per second (43 ms per item), which is similar to the encoding rate of 50 ms per item found for colours with the same memory array size by Vogel et al. (2006). The capacity of Courier New letters was approximately 2.5 items, which was lower than in Experiment 1, but nevertheless within the range reported in previous studies. Encoding rate for unfamiliar alphabets was approximately 12 items per second, and capacity was approximately 1.5 items. Estimates for Künstler, an uncommonly used decorative font, and for the Brussels Artificial Character Set, were comparable to estimates for unfamiliar letters.

Taken together, these suggest an overall effect of familiarity on the encoding rate and capacity of VWM. Xie and Zhang (2016, 2017) suggest that having preexisting long-term memory (LTM) representations of familiar stimuli enhances the speed of consolidation and capacity of VWM. Faster encoding and increased capacity were reflected in an event-related potential component known as the contralateral delay activity (CDA), which indexes the amount of encoded information in VWM (Vogel & Machizawa, 2004). CDA was significantly higher for first-generation (familiar) compared to recent-generation (unfamiliar) Pokémon characters for short encoding times, but not for longer times, consistent with a faster encoding rate. Our findings lead us to predict the same effect for familiar and unfamiliar letters, even when stimulus complexity is matched.

Letters

We found that VWM capacity was higher for familiar stimuli, but was not higher for less-complex (as assessed by perimetric complexity) stimuli. Pelli et al. (2006) reported that native readers exhibited a memory span of about four to five letters, whereas trained observers—who identified letters as efficiently as the fluent observers—could remember only about two (Pelli et al., 2006). That is, training novice observers improved identification efficiency to the level of fluent observers, but did not increase memory span.

The extensive experience of fluent readers is likely to have produced neural changes that allow neural representations of letters unavailable to untrained readers. The visual word form area (VWFA) in the left inferior occipitotemporal cortex responds selectively to visually presented words and consonant strings, and its selectivity is experience-dependent: For example, responses to Hebrew words are stronger among Hebrew readers than non-readers of Hebrew (Baker et al., 2007). This suggests that detectors in the VWFA develop templates specific to letters from overlearned alphabets. Dehaene, Cohen, Sigman and Vinckier (2005) propose a neural hierarchy of detectors in the visual system, beginning with simple feature detectors whose outputs are combined into increasingly abstract letter representations. This is consistent with functional imaging work revealing gradient of selectivity throughout the occipitotemporal cortex, through the VWFA (Vinckier et al., 2007). Repeated activation likely results in efficient pathways through this hierarchy only for familiar, overlearned characters.

Conclusion

The results of the current study indicate that the encoding rate and capacity of VWM do not vary with stimulus complexity, defined by number of basic visual features, but are markedly affected by familiarity with the memoranda. As encoding rate does not vary with complexity, we suggest that the binding of features occurs rapidly, within approximately 120 ms, after which feature-integrated objects are consolidated into VWM. The highest capacity estimates and fastest encoding rates were found for familiar English letters. This suggests that although letters are complex objects, adult expertise with them allows their representations to be activated rapidly and used for VWM consolidation. We propose that such encoding relies on an efficient neural pathway containing detectors for familiar letter shapes.

References

Alvarez, G. A., & Cavanagh, P. (2004). The Capacity of Visual Short-Term Memory Is Set Both by Visual Information Load and by Number of Objects. *Psychological Science*, *15*, 106–111.

Attneave, F., & Arnoult, M. D. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, *53*(6), 452.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual Working Memory Represents a Fixed Number of Items Regardless of Complexity. *Psychological Science*, *18*(7), 622–628. https://doi.org/10.1111/j.1467-9280.2007.01949.x

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*(3), 119–126. https://doi.org/10.1016/S1364-6613(00)01593-X

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *104*(21), 9087–9092.

Besner, D., Davies, J., & Daniels, S. (1981). Reading for meaning: The effects of concurrent articulation. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 415–437. https://doi.org/10.1080/14640748108400801

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

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

Emrich, S. M., Al-Aidroos, N., Pratt, J., & Ferber, S. (2010). Finding memory in search: The effect of visual working memory load on visual search. *The Quarterly Journal of Experimental Psychology*, *63*(8), 1457–1466. https://doi.org/10.1080/17470218.2010.483768

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, *12*(6), 1127–1133. https://doi.org/10.3758/BF03206454

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: the relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679. https://doi.org/10.3758/17.5.673

Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(4), 845.

Holcombe, A. O., & Judson, J. (2007). Visual Binding of English and Chinese Word Parts is Limited to Low Temporal Frequencies. *Perception*, *36*(1), 49–74. https://doi.org/10.1068/p5582

Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements. *Attention and Performance XVI: Information Integration in Perception and Communication*, *16*, 125–155.

Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 556.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Le-Hoa Võ, M., & Wolfe, J. M. (2015). The role of memory for visual search in scenes. *Annals of the New York Academy of Sciences*, *1339*(1), 72–81. https://doi.org/10.1111/nyas.12667

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10.1038/36846

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*(28), 4646–4674.

Vidal, C., Content, A., & Chetail, F. (2017). BACS: The Brussels Artificial Character Sets for studies in cognitive psychology and neuroscience. *Behavior Research Methods*. https://doi.org/10.3758/s13428-016-0844-8

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical Coding of Letter Strings in the Ventral Stream: Dissecting the

Inner Organization of the Visual Word-Form System. *Neuron*, *55*(1), 143–156. https://doi.org/10.1016/j.neuron.2007.05.031

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751. https://doi.org/10.1038/nature02447

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1436.

Watson, A. B. (2011, February 28). *Perimetric Complexity of Binary Digital Images: Notes on Calculation and Relation to Visual Complexity*. Retrieved from https://ntrs.nasa.gov/search.jsp?R=20110013429

Wutz, A., & Melcher, D. (2013). Temporal buffering and visual capacity: The time course of object formation underlies capacity limits in visual cognition. *Attention, Perception, & Psychophysics*, *75*(5), 921–933. https://doi.org/10.3758/s13414-013-0454-9

Xie, W., & Zhang, W. (2016). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory & Cognition*, 1–13. https://doi.org/10.3758/s13421-016-0679-7

Xie, W., & Zhang, W. (2017). Familiarity Speeds Up Visual Short-term Memory Consolidation: Electrophysiological Evidence from Contralateral Delay Activities. *Journal of Cognitive Neuroscience*, 1–13. https://doi.org/10.1162/jocn_a_01188

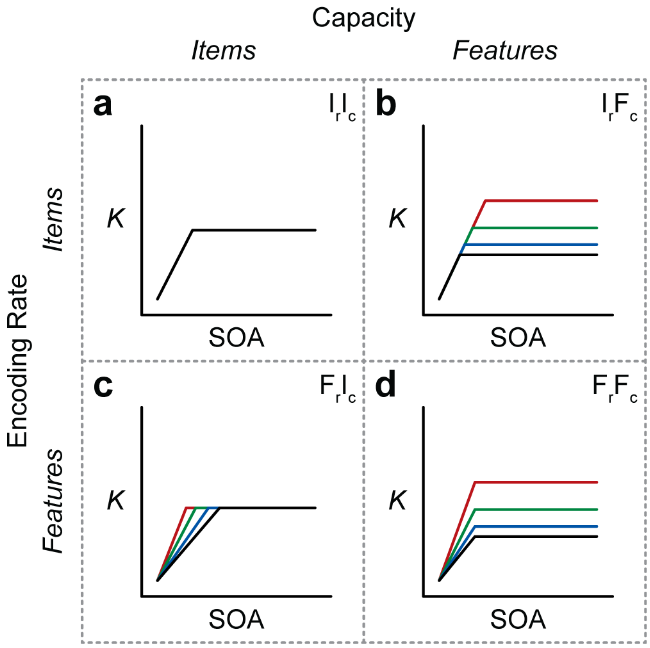

Figure 1. Storage in VWM as a function of SOA according to four models of VWM encoding. Different shades represent alphabets of different perimetric complexity. (a) R_iC_i model, by which encoding rate and capacity of VWM are quantified in terms of items, regardless of item complexity. (b) R_iC_f model, by which encoding rate is quantified in terms of items, but the capacity of VWM is defined by number of features. This results in different asymptotes for the alphabets of different complexities. (c) R_fC_i model, by which capacity of VWM is defined in terms of items, but encoding rate is quantified in terms of features. The encoding rate, reflected in the initial gradient of the function, thus varies for each alphabet according to complexity. (d) R_fC_f model, by which encoding rate and capacity of VWM are quantified in terms of features. Thus both the initial gradient and asymptote vary between alphabets according to complexity.

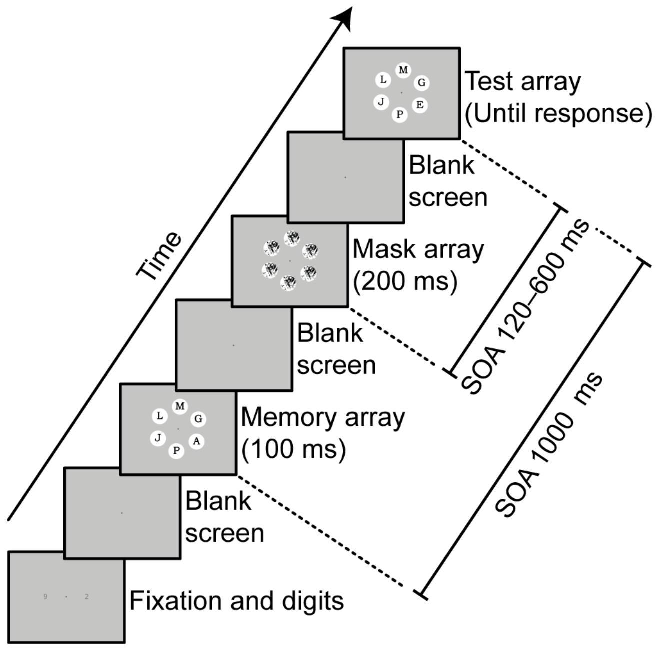

Figure 2. Stimulus sequence on a single trial. At the beginning of the trial, two digits were shown to either side of the fixation point. The memory array, containing different letters from the same alphabet, was presented for 100ms and followed by a blank screen. The dynamic mask contained phase-scrambled transformations of all letters in the alphabet, presented for 200 ms. The SOA for the memory array and mask array was 120, 130, 160, 200, 270, 390, or 600 ms on each trial. The test array, which was identical to the memory array except for a change in letter at one position, was presented 1000 ms after the memory array, and remained on the screen until the participant made a response. The participant responded to each trial by selecting the aperture in which they believed the change had occurred. For illustrative purposes, the stimuli shown here have been enlarged.

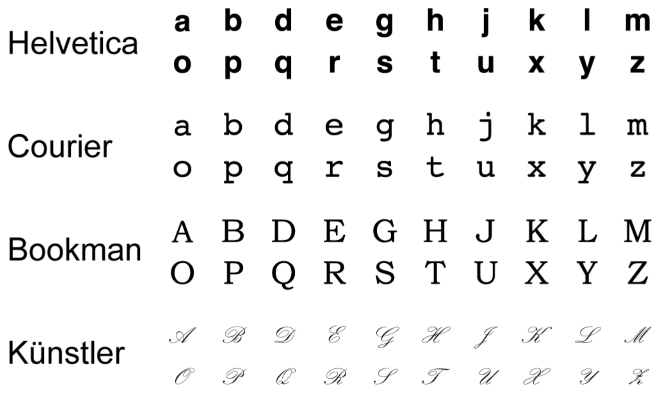

Figure 3. The twenty letter stimuli of each alphabet used in Experiment 1 (Helvetica, Courier, Bookman and Künstler).

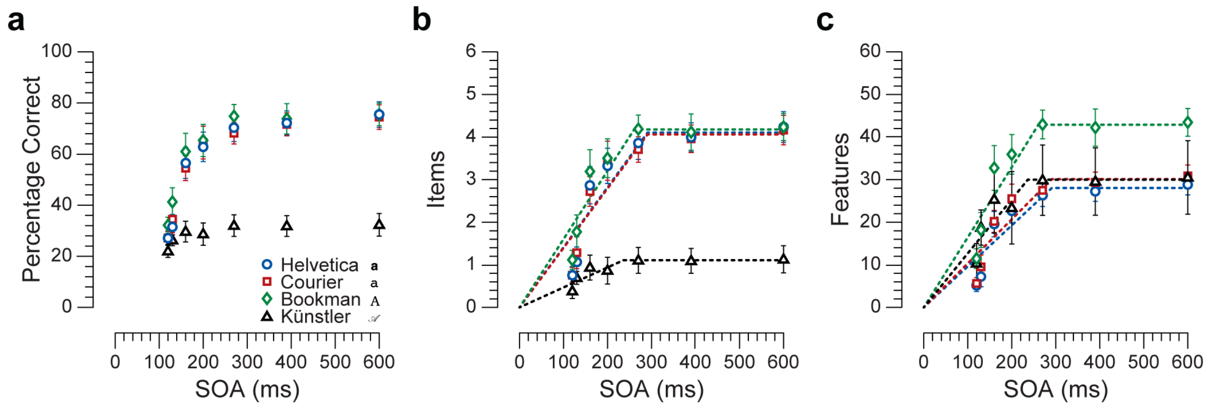

Figure 4. Results from Experiment 1. (a) Mean percentage correct on the change-detection task as a function of stimulus onset asynchrony (SOA) and alphabet. (b) Number of items encoded (K) as a function of SOA and alphabet. (c) Number of features encoded ($K\kappa$) as a function of SOA and alphabet. The number of features was estimated by multiplying the perimetric complexity (κ; see Table 1) by K. Error bars show ±1 standard error of the mean (SEM). For both (b) and (c), a two-part linear function was fitted for each alphabet; a line with a positive gradient representing the encoding rate, and a line with zero gradient at VWM capacity.

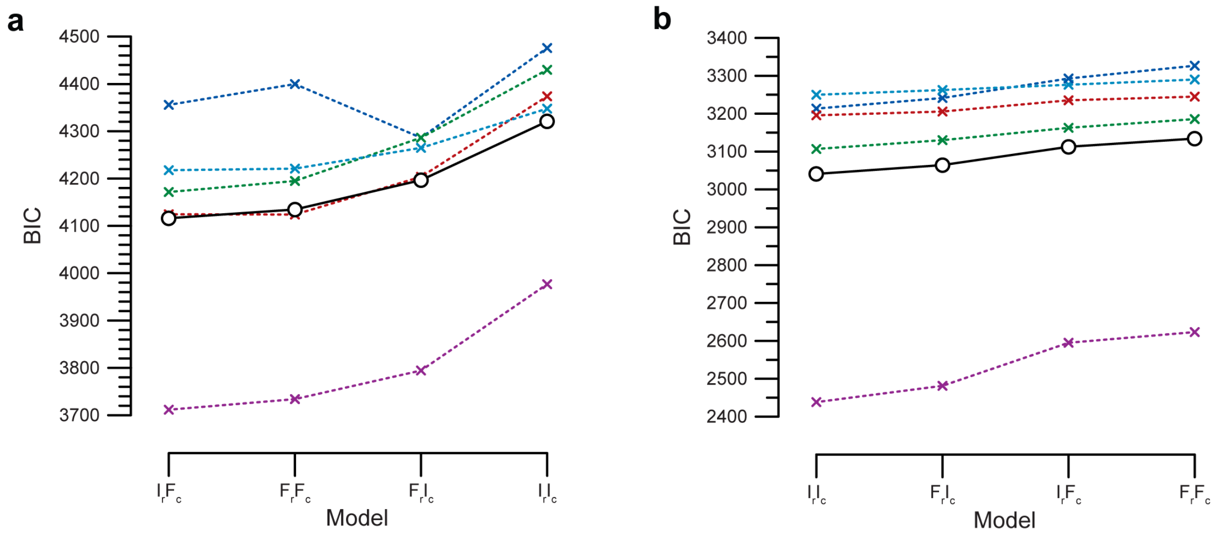

Figure 5. BIC values from maximum-likelihood model fitting in Experiment 1. (a) BIC values for each of the four candidate models including all four alphabets. BIC values for each participant are shown as crosses joined by a dashed line, and means across participants are shown as circles joined by a solid line. Models are ordered from lowest mean BIC value to greatest mean BIC value. A lower BIC value indicates a better fit. For 4 out of 5 participants, the best-fitting model was the I_rF_c model, which quantifies encoding rate in terms of items and capacity in terms of features. For the other participant, the F_rF_c model, which quantifies encoding rate and capacity in terms of features and the I_rF_c model were equally the best-fitting. (b) BIC values from maximum-likelihood model fitting in Experiment 1 with Künstler excluded. Here, the best-fitting model for all 5 participants was the I_rI_c model, which quantifies encoding rate and capacity in terms of items.

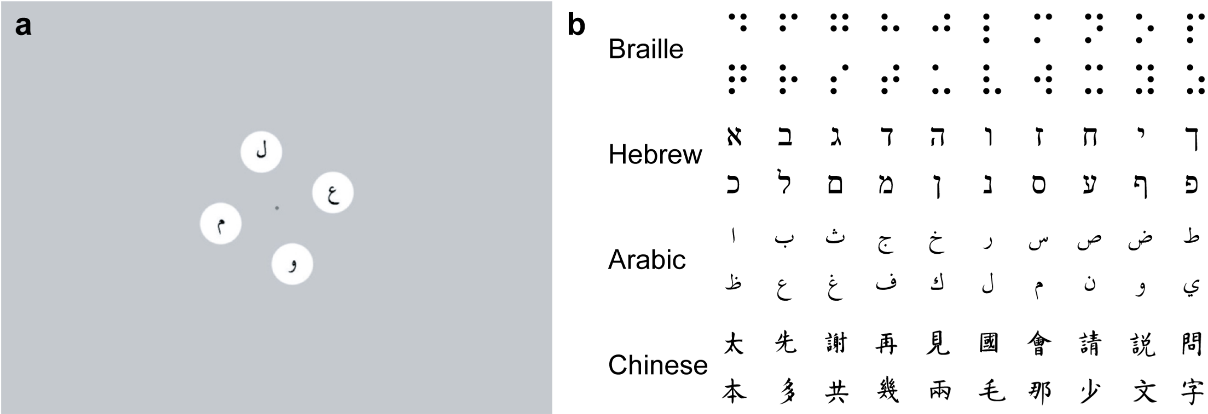

Figure 6. Stimuli for Experiment 2. (a) Example memory, showing letters from the Arabic alphabet. (b) The 20 character stimuli used for each alphabet (Braille, Hebrew, Arabic and Chinese).

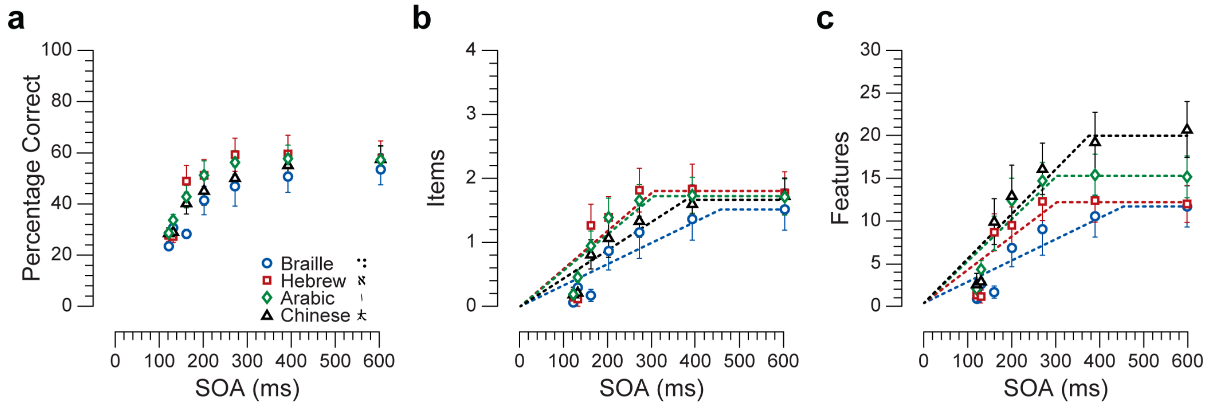

Figure 7. Results of Experiment 2. (**a**) Mean percentage correct on the change-detection task as a function of the stimulus onset asynchrony (SOA) and alphabet. (**b**) Number of items encoded (K) as a function of SOA and alphabet. (**c**) Number of features encoded ($K\kappa$) as a function of SOA and alphabet. Error bars show ±1 SEM.

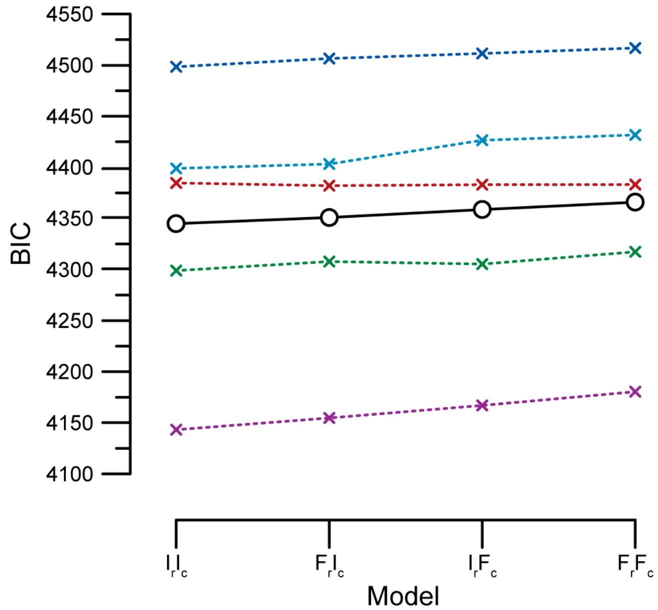

Figure 8. BIC values from maximum-likelihood model fitting in Experiment 2 for each of the four candidate models. BIC values for each participant are shown as crosses joined by a dashed line, and means across participants are shown as circles joined by a solid line. Models are ordered from lowest mean BIC value to greatest mean BIC value. For all participants, I_rI_c was the best-fitting model, which quantifies encoding rate and capacity in terms of items.

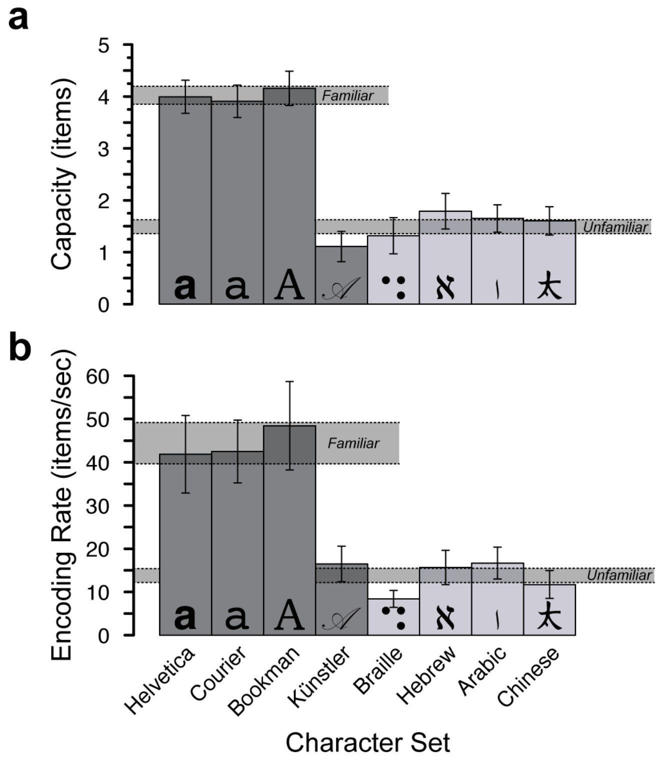

Figure 9. Mean capacity and encoding rates for all alphabets used in Experiment 1 and 2. (a) Mean capacity shows two distinctive clusters for familiar letters and novel characters. (b) Mean encoding rate likewise shows two distinct clusters. Error bars show ±1 SEM. Horizontal bars represent the range of ±1 SEM for the mean across individuals of capacity and encoding rate, separately for familiar and unfamiliar stimuli.

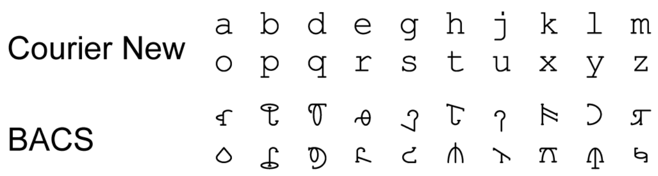

Figure 10. The twenty letter stimuli used for each alphabet in Experiment 3 (Courier New and Brussels Artificial Character Set).

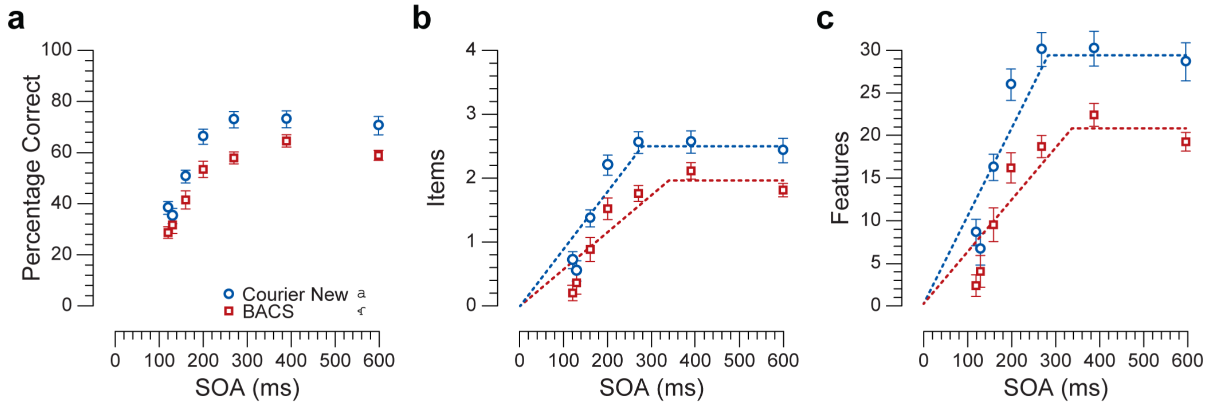

Figure 11. Results from Experiment 3. (**a**) Mean percentage correct on the change-detection task as a function of the stimulus onset asynchrony (SOA) and alphabet. (**b**) Number of items encoded (K) as a function of SOA and alphabet. (**c**) Number of features encoded ($K\kappa$) as a function of SOA and alphabet. Error bars show ±1 SEM.