The influence of learning in visual working memory

William Xiang Quan Ngiam

Contents

Chapter 1: Introduction

The visual system encounters an enormous amount of complex information that is processed to produce a smooth phenomenal experience of the world. The visual processes that achieve this remarkable feat require a memory store that encode, retain and manipulate the visual information. For example, an active memory store integrates the information between saccades (Irwin & Andrews, 1996), orients where attention should be deployed (Awh & Jonides, 2001), and retains information about objects during visual tracking and search (Carlisle, Arita, Pardo, & Woodman, 2011). The system responsible for actively storing the visual information for perception has been termed "visual working memory" (VWM). Despite its necessity in everyday perception, the VWM system is surprisingly limited in the amount of information it can encode, approximately three to four objects (Luck & Vogel, 1997). This thesis explores the processes that contribute to this capacity limit with research that examines how memory performance can be boosted to overcome this limit. This chapter provides an overview of past visual working memory research.

1.1 The conception of working memory

Classical research separated memory into two distinct but interacting systems, short-term memory (STM) and long-term memory (LTM). The STM store has a highly limited capacity that holds current information in awareness, whereas LTM is thought to be unlimited in capacity, but the information stored is effortfully retrieved (Atkinson & Shiffrin, 1968). Atkinson & Shiffrin (1968) were one of the first to consider the STM system as "working"; "a system in which decisions are made, problems are solved and information flow is directed". This early conception of STM as "working" relied on two incorrect assumption that encoding of information into LTM, and therefore learning, required maintenance in STM, which has since been shown to be untrue (Baddeley & Hitch, 1974). This was updated by Baddeley and Hitch's (1974) highly influential multi-component working memory model. Their first iteration contained three subsytems: the central executive, the phonological loop and the visuospatial sketchpad (Figure 1). The phonological loop and the visuospatial sketchpad, collectively known as the "slave systems", maintain verbal and visual information respectively.

It is the visuospatial sketchpad that is analagous to what researchers now refer to as the VWM system. The term "visual working memory" is often used synonymously with "visual short-term memory". Luck & Vogel (2013) provides three defining aspects of VWM: the information represented is visual in nature, VWM information is actively maintained and that the information is accessed for cognitive use.

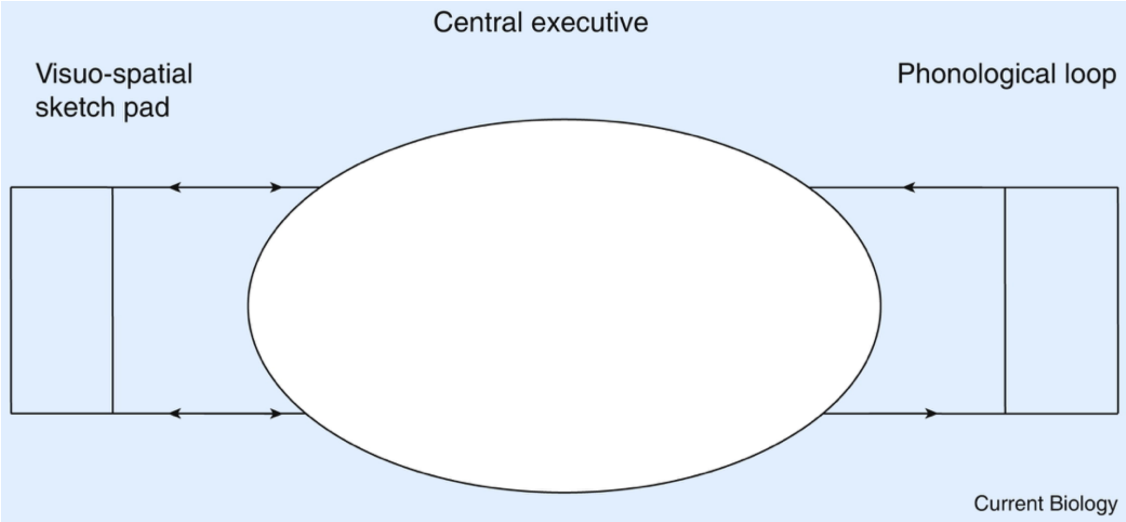

Figure 1: An early model of working memory proposed by Baddeley. Figure taken from Baddeley & Hitch (1974).

1.2 Measuring visual working memory capacity

In their seminal study, Luck & Vogel (1997) popularised the change-detection paradigm for the measurement of VWM capacity. In this paradigm (see Figure 2), an initial array (*sample array*) of objects is presented to the observer for a brief duration, usually no longer than a second, before disappearing. After a short delay, a second array (*test array*) may appear identically to the sample array (*no-change* trials) or with one object replaced by another object (*change* trials). The observer has to respond with whether they think a change occurred or not on that trial.

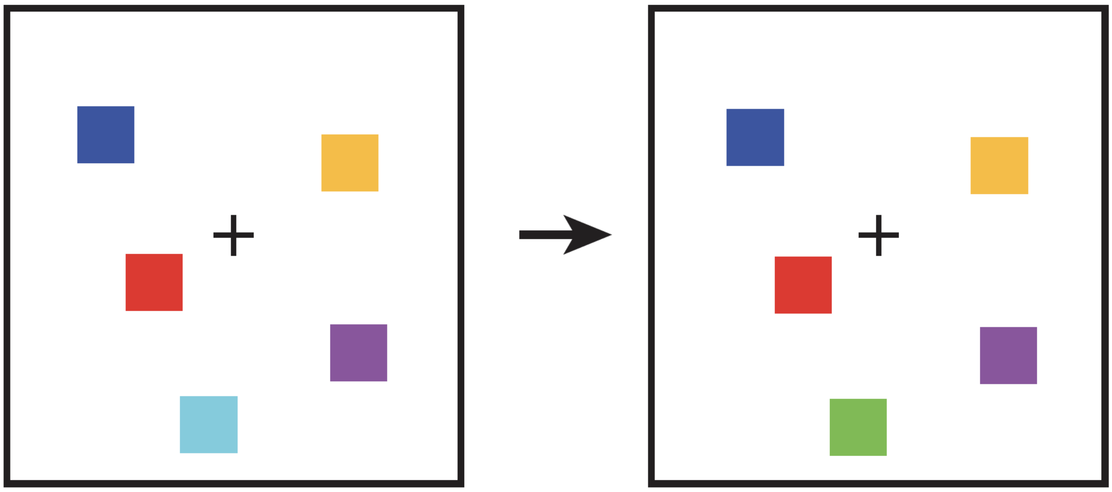

Figure 2: An example of what is displayed on a change-detection trial.

The proportion of trials that a participant correctly detects a change or no change occurred can be used to estimate the number of items held in visual working memory. Assuming the observer has stored a certain number of objects (*K*) from the sample array, a correct response on a change trial (a 'hit') will occur whenever the changed item is one of those *K* objects. If an array contains *N* objects, on average this will occur on *K* out of *N* change trials. Additional hits will occur on a porportion (*G*) of the remaining (*N*-*K*) out of *N* change trials (when the changed object is not among those encoded) if the observer correctly guesses that a change has occurred. For an unbiased observer, this will occur on half of the remaining trials (*G = 0.5*), but *G* can be estimated using the observer's false alarm rate, the overall number of trials which a change is reported when no change occurred. This produces the model proposed by Pashler (1988):

$$H = \frac{K}{N} + \frac{(N-K)}{N} * G \tag{1}$$

where *H* is the probability of a hit on a change trial. Rearranged to make *K* the subject:

$$K = \frac{N * (H - G)}{1 - G} \tag{2}$$

However, this equation assumes VWM has no bearing on a no-change trial (Cowan et al., 2005). On no-change trials, the guesswork is limited to items not stored in VWM (*N*-*K*). Cowan estimates that the subject will that a change has not occurred with a probability of *1 - G*, where *G* is the probability of guessing a change had occurred. This was updated by Cowan (2001) to include the correct rejection rate (*CR*):

$$CR = \frac{K}{N} + \frac{(N-K)}{N} * (1 - G) \tag{3}$$

Adding this to Equation 1:

$$H + CR = \frac{2K}{N} + \frac{(N-K)}{N} = \frac{(K+N)}{N} \tag{4}$$

Rearranging to make *K* the subject:

$$K = N * (H + CR - 1) \tag{5}$$

1.3 The capacity of visual working memory

Despite its necessity, the capacity of visual working memory is surprisingly limited to approximately 3-4 items' worth of information. Luck & Vogel (1997) presented sample arrays containing from 1 to 12 coloured squares for 100ms, before showing the test array approximately a second later. They found performance was almost perfect for arrays of 1 to 3 colour blocks and declined from 4 to 12 colour blocks. This pattern remained when observers were given two digits to rehearse aloud to suppress any influence of verbal working memory (see Figure 3a), when the sample duration was displayed for a longer duration, and when observers were required to only make a decision about a single cued item in the array (see Figure 3b). Estimating VWM capacity from the change-detection accuracy (see Equation 2) indicated observers stored approximately four items in VWM.

Despite agreement of this capacity limit for simple visual objects, there has been contention over the architecture of VWM producing this limit. In addition to simple colours, Luck & Vogel (1997) increased the number of relevant features in the visual stimuli presented in the same change-detection task and found an identical pattern of memory performance when presenting colours. For example, with conjunctions of colour and orientation, VWM performance was no different when instructed to detect changes in only colour, only orientation or in either feature (see Figure 3c). This pattern was also replicated with stimuli that were conjunctions of four features: colour, orientation, size and the presence of a gap (see Figure 3d) and conjunctions of the same feature type, such as two colours (see Figure 3e). Since increasing the number of relevant features in the visual stimuli did not influence memory performance, Luck & Vogel (1997) proposed that the architecture of VWM is 3 to 4 'slots' where each slot stores a representation of the visual object with its features integrated, rather than the individual features of the object.

The 'slots' model was directly opposed by the findings of Alvarez & Cavanagh (2004). In their study, participants completed the same change-detection task as in Luck & Vogel (1997) but with different stimulus sets. The stimuli sets included colour squares as Luck & Vogel (1997) had done, but also Snodgrass line drawings, shaded cubes, random polygons, Chinese characters and English letters (see Figure 4). VWM capacities were significantly different for the stimulus sets contradicting what would be predicted by the 'slots' model.

Critically, Alvarez & Cavanagh (2004) indexed the *complexity* of each stimulus set by conducting a visual search task with the same stimulus sets. In the visual search task, observers were presented a target object before asking to locate whether that target was present in an array of objects from the same stimulus set. The arrays contained either 4, 8, or 12 objects and included the target object on half the trials. The *visual search rate*, their measure of stimulus complexity, was the estimated amount of additional reaction time taken to respond that the target was present with each additional item in the array. Estimating capacity as the number of objects for each stimulus set that would correspond to 75% accuracy on the change-detection task, visual search rate was very strongly correlated (r = .992) to the inverse of capacity (see **??**).

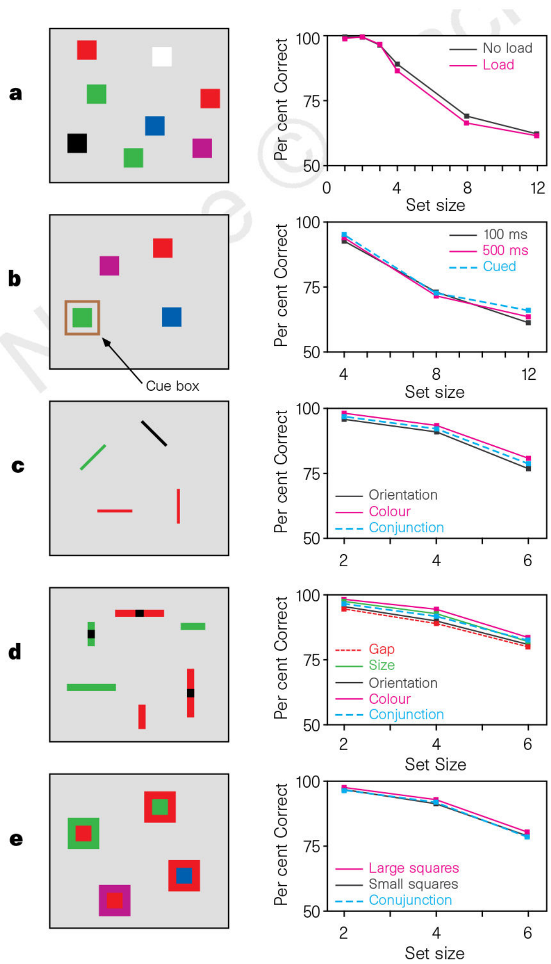

Figure 3: Stimulus arrays and memory performance from multiple experiments in Luck & Vogel (1997).

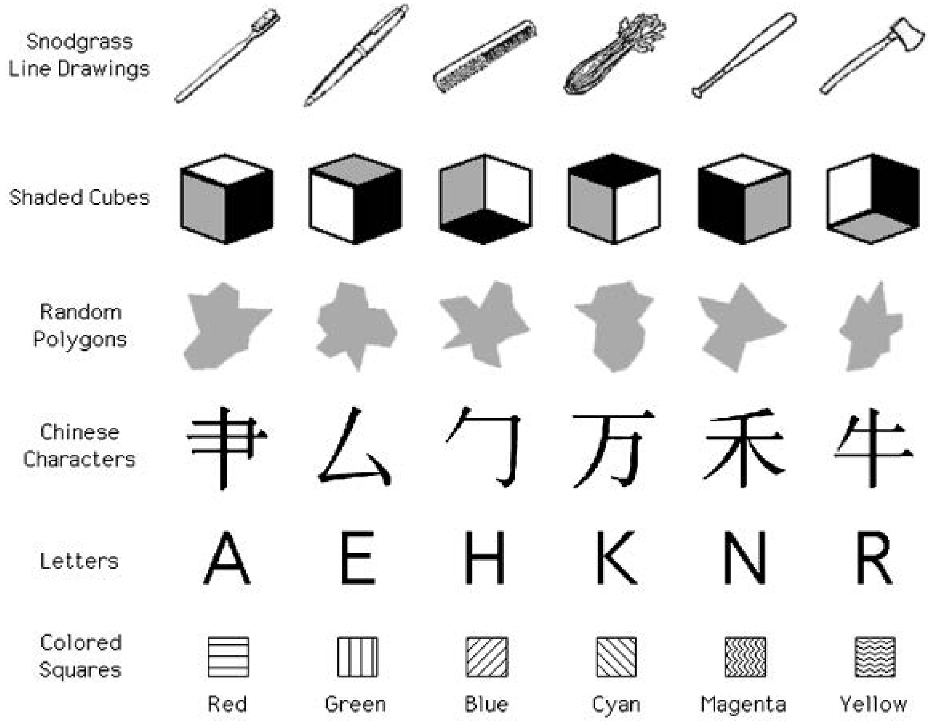

Figure 4: The stimuli sets used in Alvarez and Cavanagh (2004)

Luck & Vogel (1997) and Alvarez & Cavanagh (2004) provide contrasting findings. While Luck & Vogel (1997) show VWM capacity is consistently approximately 4 objects when varying the number of features being combined, whereas (**???**)(**???**) find VWM capacity is different for stimuli of varying complexity. Alvarez & Cavanagh (2004) suggested VWM capacity is limited by total amount of visual information rather than the number of objects as Luck & Vogel (1997) suggested in their 'slots' model. They posited the 'resource' model, which suggests that more complex visual items (those with more features) require more resources to be encoded and stored. Thus, as the visual stimuli get more complex, less items are maintained in VWM.

Awh, Barton, & Vogel (2007) disputes whether the varying VWM capacities found by Alvarez & Cavanagh (2004) was due to stimulus complexity. They suggest that the variation in VWM performance is due to an increase in comparison errors made when the object stored in memory is visually similar to the object that changed in the test array. To examine this, Awh et al. (2007) gave participants a change-detection task with memory arrays containing 4 or 8 items selected from a stimulus set of 6 shaded cubes and 6 Chinese characters. This meant that either a *within-category* change would occur, where a shaded cube changed to another shaded cube or a Chinese character changed to another Chinese character, or a *cross-category* change would occur, where a shaded cube would change to a Chinese chracter or a Chinese character would change to a shaded cube. A *within-category*

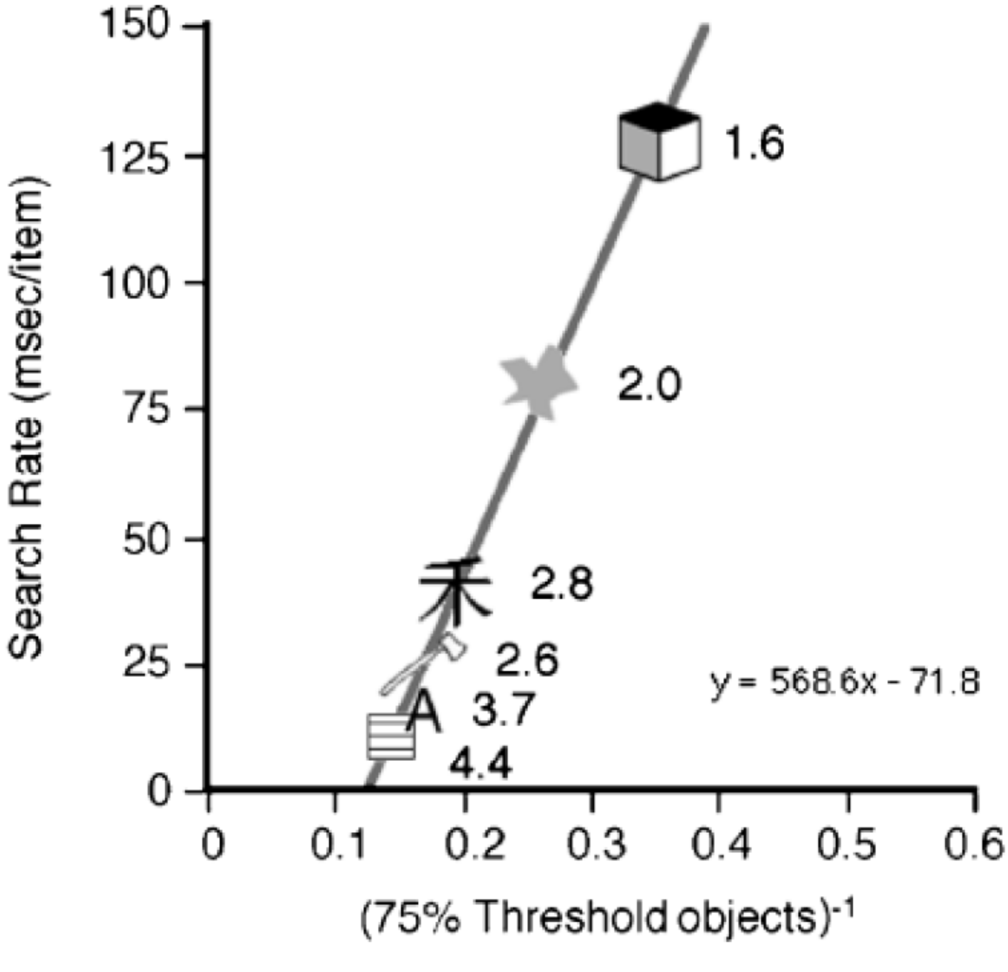

Figure 5: The visual search rate is highly correlated to the number of objects that coresop-nds to 75% accuracy on the change-detection task. The values beside each stimulus item is the calculated capacity for each stimulus set. Taken from Alvarez and Cavanagh (2004)

change is more likely to produce more errors as the to-be compared items come from the same stimulus set, whereas the to-be compared items in a *cross-category* change come from the other stimulus set and are therefore, relatively dissimilar. If stimulus complexity influences the number of items stored in VWM, then there should be no benefit of a *cross-category* change compared to a *within-category* change. But if stimulus complexity influences the comparison decision, there should be an improvement in performance for *cross-category* changes relative to *within-category* changes. They found that memory performance for *within-category* changes was significantly worse, although significantly worse for a Chinese character compared to a shaded cube. Additionally, change-detection performance for *cross-category* was equivalent for change-detection performance with colours, which has relatively low *sample-test similarity*.

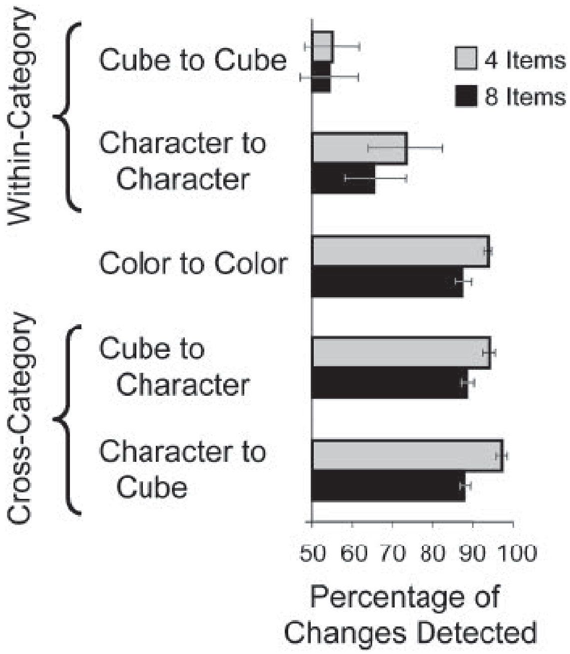

Figure 6: Results of Experiment 2 from Awh et al. (2007), showing that change-detection accuracy was significantly better for cross-category changes compared to within-category changes.

From this, Awh et al. (2007) concluded that the number of items represented in visual working memory is fixed, regardless of the complexity of those items.

1.4 Current models of visual working memory

However, the effect of training participants to be familiar with stimuli on visual working memory performance is unclear. To train recognition to polygons, Chen, Eng and Jiang (2006) presented four polygons out of a training set of eight, before presenting two polygons, one the same and one from the unpresented set. Despite being able to recognise the trained polygon, this familiarity did not improve visual working memory performance for the trained polygons over novel polygons. However, Blalock (2015) found a positive effect of familiarity training on visual working memory performance. Blalock (2015) presented a target polygon before asking the participant to select the target out of an array of four polygons. This recognition training produced better change-detection performance for trained polygons over the novel polygons. Another notable discrepancy between these studies is the sample size. While Chen et al. (2006) used twelve participants in each of their experiments, Blalock (2015) used over seventy and 102 in each of theirs. This difference in the statistical power of experiments may explain the contrasting results of familiarity training.

Chapter 2: The effect of complexity and familiarity on visual working memory

2.1 Abstract

2.2 Introduction

A common method to research the architecture of visual working memory (VWM) is to vary the stimuli used in the change-detection task on a single dimension and examining its effect on memory performance. For example, Luck & Vogel (1997) found no change in memory performance when they increased the amount of features that made up the stimuli in the change-detection task, suggesting the VWM system is object-based rather than feature-based. Similarly, in their canonical study, Alvarez & Cavanagh (2004) displayed stimuli of various complexities, such as the more complex random polygons and Chinese characters and the less complex colour squares. Critically, they indexed each stimuli's complexity by conducting a visual search task using those stimuli, in which the observer had to find a target amongst an array. The larger the visual search rate, the additional time it took to find the target with each additional item in the search display, the more complex the object was. Alvarez & Cavanagh (2004) found that change-detection performance was lower for more complex objects, such that the visual search rate, their measure of stimulus complexity, was almost perfectly correlated with working memory capacity (r = 0.992). This finding that stimulus complexity influenced memory performance motivated Alvarez & Cavanagh (2004) to suggest that the VWM system allocates finite resources to storing different stimuli, with more complex items requiring more resources.

2.2.1 Object complexity

Although the object-based "slots" model (Luck & Vogel, 1997) and the feature-based "resources" model (Alvarez & Cavanagh, 2004) have been influential in VWM research, the manner in which object complexity influences VWM processes, a main difference between these models, is still contended. Eng, Chen, & Jiang (2005) found that the visual search rates are better predictors of VWM capacities at shorter presentation durations compared to longer durations. This suggests that stimulus complexity influences perceptual encoding rather than overall VWM capacity. Awh et al. (2007) suggest that the differences in VWM capacity found by Alvarez & Cavanagh (2004) were not due to sitmulus complexity *per se* but rather because of confusion at the comparison stage in change-detection rather than during encoding. Awh et al. (2007) manipulated whether the changed object in the test array was from the same stimulus set (*within-category*) or from a different stimulus set (*cross-category*). They replicated the finding of Alvarez & Cavanagh (2004) that change-detection accuracy decreased as complexity increased with within-category changes, but found accuracy was equivalent when changes were cross-category. As objects that were more complex were more visuall similar (high *sample-test similarity*), within-category changes produced more errors made when detecting changes in the test array.

While many researchers have focused on the capacity of visual working memory, the encoding rate of information into VWM also seems to be limited. Vogel, Woodman, & Luck (2006) gradually increased the *stimulus onset asynchrony* (SOA) between the memory array and a backward-mask array in the change-detection task. They found change-detection performance improved with increased encoding duration up to 200 ms, before plateauing. Prior to the asymptote, each colour block too approximately 50 ms to encode. We used this paradigm with stimulus sets of various complexity to examine whether the encoding rate is influenced by stimulus complexity, such that an object with more features takes longer to encode into VWM, as suggested by Eng et al. (2005). Increasing object complexity may slow the rate of encoding into VWM, such that complex objects will require more time to fill VWM capacity. This would confound conclusions made from comparisons of VWM capacity for objects of different complexity with the same memory array durations, such as those found by Alvarez & Cavanagh (2004).

Differing definitions and measures of complexity may have lead to the vastly differing models of VWM architecture. Here, we defined stimulus complexity using *perimetric complexity*, the square of the combined inside and outside perimeters of a letter, divide by its area (Attneave & Arnoult, 1956). As letters increase in perimetric complexity, they are identified increasing inefficiently (Pelli, Burns, Farell, & Moore-Page, 2006). Perimetric complexity is a superior measure of complexity to previous manipulations of complexity because it is objective, measured from the stimulus directly. Additionally, an increase in perimetric complexity reflects an increase in stimulus complexity without the addition of extra feature dimensions. In the present study, we selected letters of the English alphabet and varied the perimetric complexity by presenting the letters in four different fonts (Experiment 1), as well as presented characters from four alphabets that were unfamiliar to our participants (Experiment 2).

2.3 Encoding rate

An additional factor that has been shown to influence consolidation and storage in VWM is familiarity. For example, chess experts showed an improved memory performance for chess game positions compared to novices, but equivalent memory performance when the chess pieces were random on the board (Chase & Simon, 1973). More recently, higher VWM capacities have been found for famous faces over unfamiliar faces (Jackson & Raymond, 2008), as well as for Pokémon (characters from a popular childhood cartoon) from an original generation over a recent generation only for those reporting familiarity with the characters (Xie & Zhang, 2016). Similarly, those familiar showed a higher encoding rate for Pokémon (Xie & Zhang, 2017). Although, these studies do not control stimulus complexity and it is unknown whether these effects of familiarity are independent of stimulus complexity. We examined this in Experiment 3, controlling for stimulus complexity by using the Brussels Artificial Character Set (BACS) (Vidal, Content, & Chetail, 2017). The BACS is designed to have the same number of junctions, strokes and terminations as English letters but is unfamiliar to the observer. Additionally, we matched the perimetric complexity of the BACS to the English letters.

Encoding and capacity limits in VWM might best be described in terms of *objects*, as in the "slots" model or in terms of *features*, as in the "resources" model. If feature intergration limits the encoding rate into VWM, more complex letters will be encoded at a slower rate. If this is not the case, encoding rate will not vary with stimulus complexity. Similarly, if the number of features limits VWM capacity, fewer items will be stored from more complex alphabets. Otherwise, VWM capacity maybe be determined by the number of items, and will not vary with stimulus complexity. These models are shown in Figure 1.

Chapter 3: The effect of statistical learning on visual working memory

3.1 Abstract

In the previous chapter, the effect of familiarity on VWM was examined…

3.2 Introduction

Given the strict capacity limit of VWM and its robust correlations to measures of cognitive ability such as fluid intelligence and scholastic achievement (A. R. A. Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Cowan, Chen, & Rouder, 2004; Fukuda, Vogel, Mayr, & Awh, 2010; Unsworth, Fukuda, Awh, & Vogel, 2014), there is sustained interest in manipulations that potentially enhance VWM capacity. An influential paper by Brady, Konkle, & Alvarez (2009) demonstrated including regularities in the color pairs shown in displays improved VWM recall performance relative to displays without any regularities. That is, when specific pairs of colours were more likely to appear in the display, observers were able to use this to their advantage and improve recall accuracy. This chapter explores This set of experiments investigated the mechanism of learning that produces the enhanced memory performance, and how that maps on to the architecture of VWM.

3.2.1 Chunking

Improvement in memory performance has classically been explained using *chunking*. Miller (1956) proposed learning allowed greater amounts of information to be stored more efficiently in "chunks", with the absolute number of chunks stored into memory remaining constant (Miller, 1956). For example, experts recall chess positions from real matches significantly better than novices, but not better when these positions are random distributed chess pieces (**???**). As recall performance was equivalent on random positions, this suggests chess masters do not have a larger VWM capacity but instead use their expertise to efficiently "chunk" game positions, which novices cannot do.

Brady et al. (2009) show chunking models provide an accurate approximation of how observers may take advantange of the statistical regularities.

List of Figures

List of Tables

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. Retrieved from http://visionlab.harvard.edu/Members/Patrick/PDF.files/2005%20pdfs/Alvarez%26Cavanagh(2004).pdf

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes. In (Vol. 2, pp. 89–195). Elsevier. Retrieved from http://dx.doi.org/10.1016/s0079-7421(08)60422-3

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

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

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. doi:10.1111/j.1467-9280.2007.01949.x

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press. Retrieved from http://www.sciencedirect.com/science/article/pii/S0079742108604521

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in Visual Working Memory: Using Statistical Regularities to Form More Efficient Memory Representations. *Journal of Experimental Psychology: General*, *138*(4), 487–502. Retrieved from https://www.researchgate.net/profile/Talia_Konkle/publication/38061305_Compression_in_visual_working_memory_using_statistical_regularities_to_form_more_efficient_memory_representations/links/09e4150c87249c1627000000.pdf

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional Templates in Visual Working Memory. *Journal of Neuroscience*, *31*(25), 9315–9322. doi:10.1523/JNEUROSCI.1097-11.2011

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55–81. doi:10.1016/0010-0285(73)90004-2

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*(2), 163–183. doi:10.1016/S0160-2896(01)00096-4

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences; New York*, *24*(1), 154–176. Retrieved from https://search-proquest-com.ezproxy1.library.usyd.edu.au/docview/212310135/abstract/FFB98D1D6F504F6CPQ/1

Cowan, N., Chen, Z., & Rouder, J. N. (2004). Constant Capacity in an Immediate Serial-Recall Task: A Logical Sequel to Miller (1956). *Psychological Science*, *15*(9), 634–640. doi:10.1111/j.0956-7976.2004.00732.x

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. doi:10.1016/j.cogpsych.2004.12.001

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, *12*(6), 1127–1133. doi: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. doi:10.3758/17.5.673

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. Retrieved from https://books.google.com/books?hl=en&lr=&id=HUcGqC-O8i0C&oi=fnd&pg=PA125&ots=UWQsVUeSfh&sig=R6Df_G3AVOHh__q_XjHHILzVxvA

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.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. doi: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. doi:10.1016/j.tics.2013.06.006

Miller, G. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review*. Journal (Paginated).

Retrieved from http://cogprints.org/730/

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), 369–378. doi:10.3758/BF03210419

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

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26. doi:10.1016/j.cogpsych.2014.01.003

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

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), 1436–1451. doi:10.1037/0096-1523.32.6.1436

Xie, W., & Zhang, W. (2016). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory & Cognition*, 1–13. doi: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. doi:10.1162/jocn_a_01188