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_integration_1996], orients where attention should be deployed [@awh_overlapping_2001], and retains information about objects during visual tracking and search [@carlisle_attentional_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_capacity_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.

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_human_1968]. @atkinson_human_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_working_1974]. This was updated by Baddeley and Hitch’s [-@baddeley_working_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.

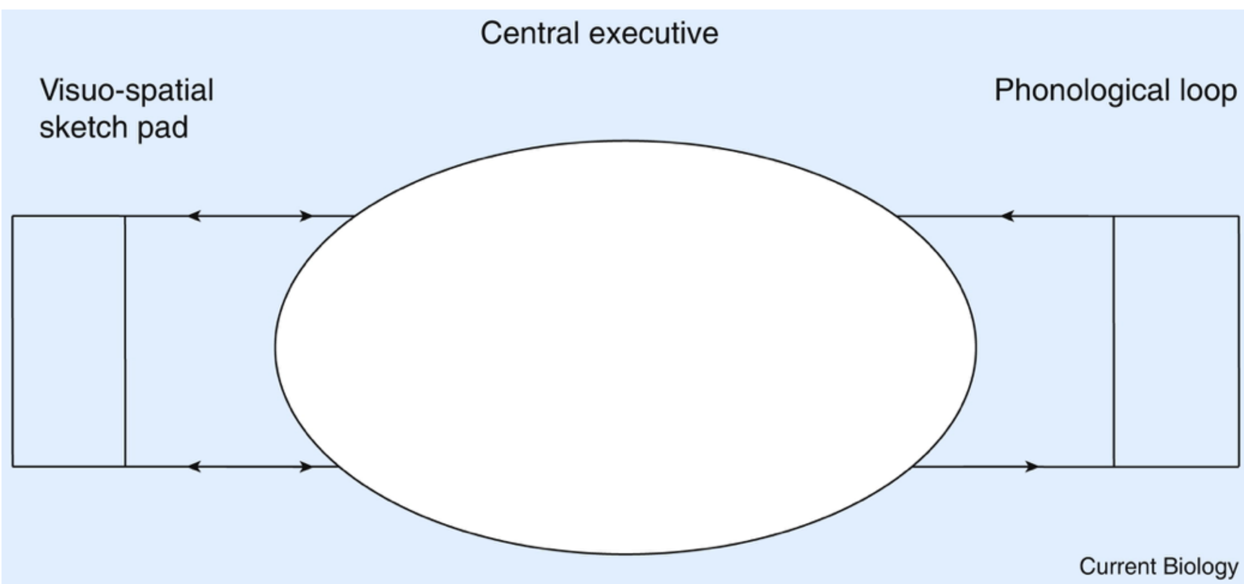

Figure 1: An early model of working memory proposed by Baddeley. Figure taken from @baddeley_working_1974.

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_visual_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.

Measuring visual working memory capacity

In their seminal study, @luck_capacity_1997 popularised the change-detection paradigm for the measurement of VWM capacity.

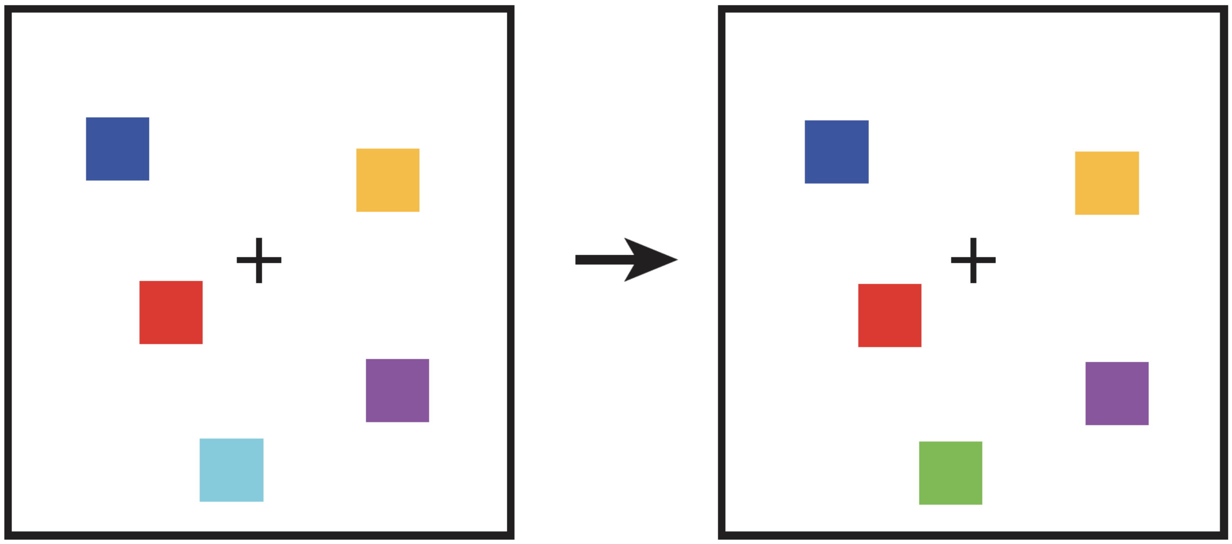

Figure 2: An example of what is displayed on a change-detection trial.

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.

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_familiarity_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_capacity_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_metatheory_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:

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

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_capacity_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_capacity_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_capacity_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 first contested by the findings of @alvarez_capacity_2004. In their study, participants completed the same change-detection task as in @luck_capacity_1997 but with different stimulus sets, such as random polygons and chinese characters (see *Figure 2*), finding different VWM capacities for different stimulus sets. Critically, they also indexed the complexity of each stimulus set by measuring the amount of additional time taken to find the target in a visual search task with each added item to the visual search array. This visual search rate, their measure of complexity, was strongly correlated with the VWM capacity, such that the more complex the item was, the lower the capacity.

Since VWM capacities varied across the different stimulus sets, and strongly correlated with stimulus complexity, @alvarez_capacity_2004 suggested VWM capacity varied with the amount of features contained in the stimuli. They posited the 'resource' model, which suggests that visual items with more features require more resources to be encoded and stored. Thus, as the visual stimuli get more complex, less items can be stored.

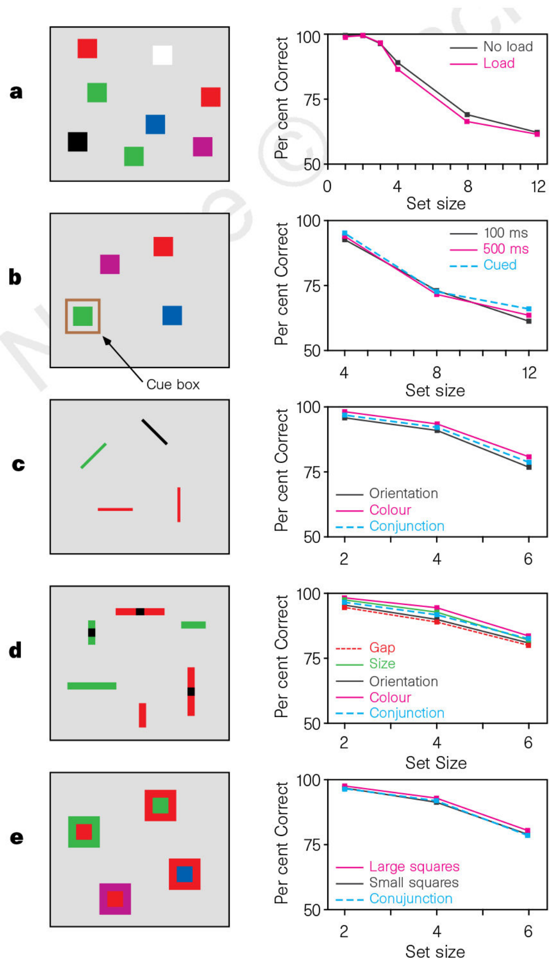

Figure 3: Stimulus arrays and memory performance from multiple experiments in @luck_capacity_1997.

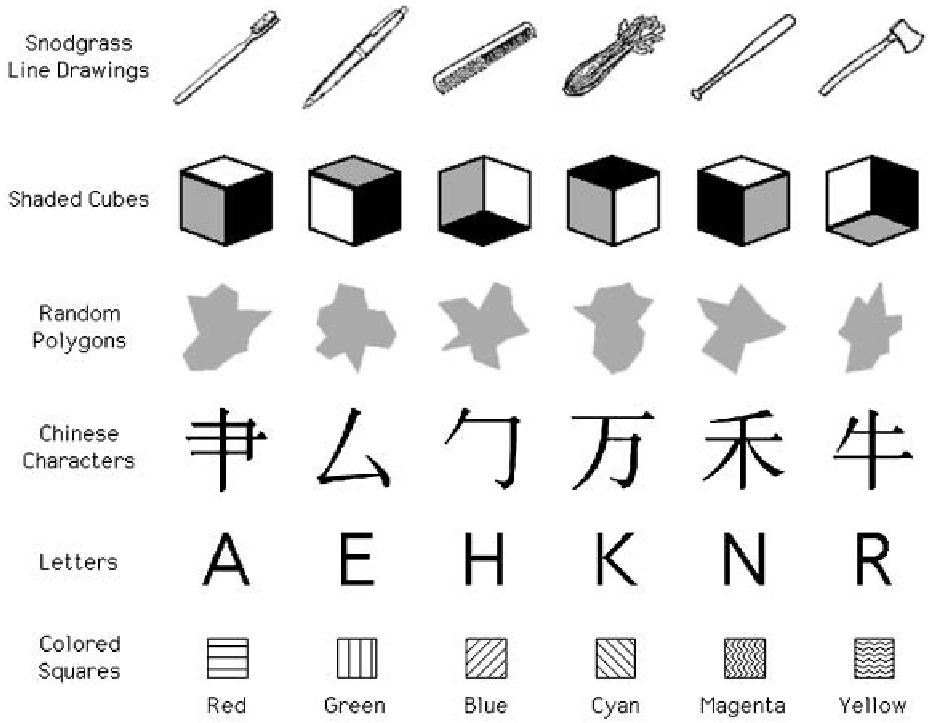

Figure 4: The stimuli sets used in Alvarez and Cavanagh (2004)

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.