Cognitive Psychology 89 (2016) 71–105

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

The attention-weighted sample-size model of visual short-term memory: Attention capture predicts resource allocation and memory load

Philip L. Smith [a,*], Simon D. Lilburn [a], Elaine A. Corbett [a], David K. Sewell [a], Søren Kyllingsbæk [b]

[a] Melbourne School of Psychological Sciences, The University of Melbourne, Australia
[b] Department of Psychology, The University of Copenhagen, Copenhagen, Denmark

ARTICLE INFO

Article history:
Accepted 13 July 2016
Available online 2 August 2016

Keywords:
Short-term memory
Working memory
Capacity limitations
Sample-size model
Diffusion model

ABSTRACT

We investigated the capacity of visual short-term memory (VSTM) in a phase discrimination task that required judgments about the configural relations between pairs of black and white features. Sewell et al. (2014) previously showed that VSTM capacity in an orientation discrimination task was well described by a sample-size model, which views VSTM as a resource comprised of a finite number of noisy stimulus samples. The model predicts the invariance of $\sum_i (d'_i)^2$, the sum of squared sensitivities across items, for displays of different sizes. For phase discrimination, the set-size effect significantly exceeded that predicted by the sample-size model for both simultaneously and sequentially presented stimuli. Instead, the set-size effect and the serial position curves with sequential presentation were predicted by an attention-weighted version of the sample-size model, which assumes that one of the items in the display captures attention and receives a disproportionate share of resources. The choice probabilities and response time distributions from the task were well described by a diffusion decision model in which the drift rates embodied the assumptions of the attention-weighted sample-size model.

* Corresponding author at: Melbourne School of Psychological Sciences, The University of Melbourne, VIC 3010, Australia.
E-mail address: philipls@unimelb.edu.au (P.L. Smith).

http://dx.doi.org/10.1016/j.cogpsych.2016.07.002
0010-0285/

1. Introduction

Visual short-term memory (VSTM), or working memory, has been identified as one of the primary bottlenecks or sources of capacity limitation in simple cognitive tasks, particularly in those tasks requiring decisions about briefly presented stimuli. Because of VSTM's theoretical importance as a source of capacity limitations, researchers have devoted considerable effort to attempting to characterize the structure and function of VSTM and the way in which its properties interact with other cognitive processes, such as perception, attention, and decision-making. Recent theoretical and experimental work has focused on whether VSTM capacity is best characterized as an item (or "slot") capacity limitation (Awh, Barton, & Vogel, 2007; Rouder et al., 2008; Shibuya & Bundesen, 1988), a feature capacity limitation (Fougnie & Alvarez, 2011), a resource capacity limitation (Bays, Catalao, & Husain, 2009; Vogel, Woodman, & Luck, 2001), or some combination of these (Alvarez & Cavanagh, 2004; Donkin, Nosofsky, Gold, & Shiffrin, 2013; van den Berg, Awh, & Ma, 2014). A variety of different experimental methods have been used to investigate how stimulus representations in VSTM are affected by memory load, including change detection (Kyllingsbæk & Bundesen, 2009; Pashler, 1988; Vogel, Woodman, & Luck, 2006), two-alternative forced choice (2AFC) discrimination (Pearson, Raškevičius, Bays, Pertzov, & Husain, 2014; Sewell, Lilburn, & Smith, 2014), confidence ratings (Donkin, Tran, & Nosofsky, 2014; Rouder et al., 2008), and continuous report (Wilken & Ma, 2004; Zhang & Luck, 2008).

Sewell et al. (2014) investigated VSTM for Gabor patch stimuli (Gaussian vignetted sinusoidal gratings) in small (one to four item) displays using an orientation discrimination task, and found that memory for these stimuli was well described by a *sample-size* model (Bonnel & Hafter, 1998; Bonnel & Miller, 1994; Heath, 1972;Lappin & Bell, 1976; Lindsay, Taylor, & Forbes, 1968; Palmer, 1990; Shaw, 1980; Swets, Shipley, McKey, & Green, 1959; Taylor, Lindsay, & Forbes, 1967). The sample-size model views VSTM as a resource, comprised of a set of independent, noisy, evidence samples that are available to represent the stimuli. If samples are recruited at a constant rate during stimulus exposure, then the number of evidence samples, n, will be proportional to exposure duration. When there is only one stimulus in the display, all n samples are available to represent it. When there are m stimuli in the display, and if there is no preferential weighting of items by attention, then each stimulus will be represented by n/m samples.

The signature prediction of the sample-size model is the invariance of $\sum_i (d'_i)^2$, where d' is the sensitivity measure of signal detection theory. If d'_m is the sensitivity for a single item when there are m items in the display, then the model predicts that

$$d'_m = \frac{d'_1}{\sqrt{m}}, \quad m = 2, 3, \ldots \tag{1}$$

or equivalently,

$$\sum_{i=1}^{m} (d'_i)^2 = c, \quad \text{(constant)}. \tag{2}$$

Eq. (2) states that the sum of squared item sensitivities will be the same for displays of different sizes, m.

These predictions follow from elementary sampling theory. The model assumes that stimulus discriminability depends on the sum (or equivalently, the mean) of the sample values that represent it. The expected value of the sum and the variance of the sum will both be proportional to n/m. Signal detection d' is a measure of signal-to-noise ratio, which depends on the ratio of the mean to the standard deviation. For a statistic based on the mean of n samples, the standard deviation is the standard error of the mean, and the ratio of the mean and standard deviation is proportional to $\sqrt{n/m}$. For fixed n, d' will be inversely proportional to $\sqrt{m}$. Eqs. (1) and (2) follow from this fact.

A striking feature of the sample-size model is that its predictions are entirely parameter-free, at least to a first approximation. For a given exposure duration, Eq. (1) predicts performance on displays