Cognitive Psychology 97 (2017) 79–97

Contents lists available at ScienceDirect

Cognitive Psychology

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

Clear evidence for item limits in visual working memory

Kirsten C.S. Adam [a,b,*], Edward K. Vogel [a,b,c], Edward Awh [a,b,c,*]

[a] Institute for Mind and Biology, University of Chicago, United States
[b] Department of Psychology, University of Chicago, United States
[c] Grossman Institute for Neuroscience, Quantitative Biology, and Human Behavior, University of Chicago, United States

ARTICLE INFO

Article history:
Accepted 6 July 2017

Keywords:
Visual working memory
Capacity limits
Precision
Metacognition

ABSTRACT

There is a consensus that visual working memory (WM) resources are sharply limited, but debate persists regarding the simple question of whether there is a limit to the total number of items that can be stored concurrently. Zhang and Luck (2008) advanced this debate with an analytic procedure that provided strong evidence for random guessing responses, but their findings can also be described by models that deny guessing while asserting a high prevalence of low precision memories. Here, we used a whole report memory procedure in which subjects reported all items in each trial and indicated whether they were guessing with each response. Critically, this procedure allowed us to measure memory performance for all items in each trial. When subjects were asked to remember 6 items, the response error distributions for about 3 out of the 6 items were best fit by a parameter-free guessing model (i.e. a uniform distribution). In addition, subjects' self-reports of guessing precisely tracked the guessing rate estimated with a mixture model. Control experiments determined that guessing behavior was not due to output interference, and that there was still a high prevalence of guessing when subjects were instructed not to guess. Our novel approach yielded evidence that guesses, not low-precision representations, best explain limitations in working memory. These guesses also corroborate a capacity-limited working memory system – we found evidence that subjects are able to report non-zero information for only 3–4 items. Thus, WM capacity is constrained by an item limit that precludes the storage of more than 3–4 individuated feature values.

1. Introduction

Working memory (WM) is an online memory system where information is maintained in the service of ongoing cognitive tasks. Although there is a broad consensus that WM resources are sharply limited, there has been sustained debate about the precise nature of these limits. On the one hand, discrete-resource models argue that only a handful of items can be maintained at one time, such that some items fail to be stored when the number of memoranda exceeds the observer's capacity (Awh, Barton, & Vogel, 2007; Cowan, 2001; Rouder et al., 2008; Zhang & Luck, 2008). On the other hand, continuous resource models argue that WM storage depends on a central pool of resources that can be divided across an unlimited number of items (Bays & Husain, 2008; van den Berg, Shin, Chou, George, & Ma, 2012; Wilken & Ma, 2004).

Of course, it has long been known that memory performance declines as the number of memoranda increases in a WM task. For example, Luck and Vogel (1997) varied the number of simple colors in a change detection task that required

* Corresponding authors at: University of Chicago, 940 E 57th St, Chicago, IL 60637, United States.
E-mail addresses: kadam1@uchicago.edu (K.C.S. Adam), awh@uchicago.edu (E. Awh).

http://dx.doi.org/10.1016/j.cogpsych.2017.07.001
0010-0285/

subjects to detect whether one of the memoranda had changed between two sequential presentations of a sample display. They found that while performance was near ceiling for set sizes up to three items, accuracy declined quickly as set size increased beyond that point. This empirical pattern is well described by a model in which subjects store 3–4 items in memory and then fail to store additional items. However, the same data can be accounted for by a continuous resource model that presumes storage of all items, but with declining precision as the number of memoranda increases (Wilken & Ma, 2004). According to continuous resource accounts, increased errors with larger set sizes are caused by insufficient mnemonic precision rather than by storage failures (but for a critique of this account see Nosofsky & Donkin, 2016b). Thus, a crux issue in this literature has been to distinguish whether performance declines with displays containing more than a handful of items are due to storage failures or sharp reductions in mnemonic precision.

In this context, Zhang and Luck (2008) offered a major step forward with an analytic approach that provides separate estimates of the probability of storage and the quality of the stored representations. They employed a continuous recall WM task in which subjects were cued to recall the precise color of an item from a display with varying numbers of memoranda. Their key insight was that if subjects failed to store a subset of the items, there should be two qualitatively distinct types of responses within a single distribution of response errors. If subjects had stored the probed item in memory, responses should be centered on the correct color, with a declining frequency of responses as the distance from the correct answer increased. But if subjects had failed to store the probed item, then responses should be random with respect to the correct answer, producing a uniform distribution of answers across the entire space of possible colors. Indeed, their data revealed that the aggregate response error distribution was well described as a weighted average of target-related and guessing responses. Thus, Zhang and Luck (2008) provided some of the first positive evidence that working memory performance reflects a combination of target-related and guessing responses.

Subsequent work, however, has argued that the empirical pattern reported by Zhang and Luck (2008) can be explained by continuous resource models that presume storage of all items in every display (van den Berg, Awh, & Ma, 2014; van den Berg et al., 2012). A key feature of these models has been the assumption that precision in visual WM may vary substantially; thus, while some items may be represented precisely, other representations in memory may contain little information about the target item. Using this assumption, van den Berg et al. (2012) showed that they could account for the full distribution of errors – including apparent guessing – and that their model outperformed the one proposed by Zhang and Luck (2008). Indeed, converging evidence from numerous studies has left little doubt that precision varies across items in these tasks (e.g., Fougnie, Suchow, & Alvarez, 2012). That said, the question of whether precision is variable is logically separate from the question of whether observers ever fail to store items in these procedures. To examine the specific reasons why one model might achieve a superior fit over another, it is necessary to explore how distinct modeling decisions influence the outcome of the competition. Embracing this perspective, van den Berg et al. (2014) carried out a factorial comparison of WM models in which the presence of items limits and the variability of precision were independently assessed. Although this analysis provided clear evidence that mnemonic precision varies across items and trials, the data were not decisive regarding the issue of whether working memory is subject to an item limit. There was a numerical advantage for models that endorsed item limits, but it was not large enough to draw strong conclusions. Thus, the critical question of whether item limits in visual working memory elicit guessing behavior remains unresolved.

Here, we report data that offer stronger traction regarding this fundamental question about the nature of limits in working memory. Much previous work has focused on explaining variance within aggregate response-error distributions (i.e. the shape of the response distribution and how it changes across set sizes). Here, we chose a different route. Rather than developing a new model that might explain a small amount of additional variance in "traditional" partial report datasets, we developed a new experimental paradigm in which subjects recalled—in any order that they wished—the precise color (Experiment 1a) or orientation (Experiment 1b) of every item in the display. This procedure has the key benefit of measuring the quality of all simultaneously remembered items, and it yields the clear prediction that if there are no item limits, then there should be measurable information across all responses. To anticipate the results, this whole report procedure provided rich information about the quality of all items within a given trial as well as subjects' metaknowledge of variations in quality. Observers consistently reported the most precisely remembered items first, yielding monotonic declines in information about the recalled item with each successive response. Critically, for the plurality of subjects, the final three responses made were best modeled by the parameter-free uniform distribution that indicates guessing. In additional analyses and experiments, we showed that subjective guess ratings tracked mixture model guessing parameter (Experiments 1 & 2), that output interference could not explain our estimates of capacity (Experiment 2), and that making subjective guess ratings did not drive our evidence for guessing (Experiment 3). Finally, we used simulations to question a key claim of the variable precision model – that representations used by this model all contain measurable information. Previously, others have suggested that the variable precision model may mimic guess responses with ultra-low precision representations (Nosofsky & Donkin, 2016b; Sewell, Lilburn, & Smith, 2014). Here, we advanced these claims by showing that variable precision models that eschew guessing posit a high prevalence of memories that are indistinguishable from guesses. Moreover, the frequency of these putative representations precisely tracked the estimated rate of guessing in models that acknowledge item limits.

In sum, there has been a longstanding debate over whether there is any limit in the number of items that can be stored in working memory. Our findings provide compelling evidence that working memory is indeed subject to item limits, disconfirming a range of prior models that deny guessing entirely or posit an item limit that varies from trial to trial without any hard limit in the total number of items that can be stored (e.g. Sims, Jacobs, & Knill, 2012; van den Berg et al., 2014). Instead,