Discrete fixed-resolution representations in visual working memory

Weiwei Zhang[1,2] & Steven J. Luck[2]

Limits on the storage capacity of working memory significantly affect cognitive abilities in a wide range of domains[1], but the nature of these capacity limits has been elusive[2]. Some researchers have proposed that working memory stores a limited set of discrete, fixed-resolution representations[3], whereas others have proposed that working memory consists of a pool of resources that can be allocated flexibly to provide either a small number of high-resolution representations or a large number of low-resolution representations[4]. Here we resolve this controversy by providing independent measures of capacity and resolution. We show that, when presented with more than a few simple objects, human observers store a high-resolution representation of a subset of the objects and retain no information about the others. Memory resolution varied over a narrow range that cannot be explained in terms of a general resource pool but can be well explained by a small set of discrete, fixed-resolution representations.

To separately measure the number of items stored in working memory and the precision of each representation, we used a short-term recall paradigm[5,6] in which subjects report the remembered colour of a probed item by clicking on a colour wheel (Fig. 1a). If the probed item has been stored in working memory, the recalled value will tend to be near the original colour. If the probed item has not been stored, then the observer will have no information about the colour, and the responses should be random. These two types of trials are mixed together in the data (Fig. 1b), but the components can be recovered via standard estimation methods. This produces one parameter (P_m) representing the probability that the probed item was present in memory at the time of the probe and another parameter (s.d.) representing the precision of the representation when the cued item was present in memory.

Experiment 1 ($N = 8$) tested this model using set sizes of 3 or 6 coloured squares (Fig. 1c). s.d. did not vary significantly across set sizes ($F < 1$), whereas P_m was approximately twice as great at set size 3 as at set size 6 ($F(1,7) = 761.26$, $P < 0.001$). Our simple fixed-resolution model provided an excellent quantitative fit to the data, whereas a model in which all items are encoded could not fit the data (see Supplementary Notes). This result rules out the entire class of working memory models in which all items are stored but with a resolution or noise level that depends on the number of items in memory[5]. Control experiments demonstrated that these results cannot be explained by a lack of time to encode the items or by a lack of sensitivity, and additional analyses demonstrated that the observers remembered continuous colour values rather than colour categories (see Supplementary Notes).

These results demonstrate that observers store a small number of representations with good precision. However, it is possible that performance is influenced both by a limited number of 'storage slots' and a limited pool of resources[7]. As an analogy, consider three cups (the slots) and a bottle of juice (the resource). It would be impossible

to serve juice to more than three people at a time, but it would be possible to pour most of the juice into a single cup, leaving only a few drops for the other two cups. Thus, allocating most of the resources to a single representation could increase the precision of that representation, leaving 'only a few drops' of resources for the other representations, which would then be highly imprecise. We call this the 'slots+resources' model.

The storage of information in visual working memory could instead be an all-or-none process that either creates a representation of a given precision or creates no representation at all. This would be analogous to a limited set of prepackaged juice boxes of a fixed size. The juice boxes are still a type of resource, but one that is highly constrained by the small number and fixed size of each box. That is, if three juice boxes are available, an individual could be given 0, 1, 2 or 3 boxes. Similarly, if three memory slots are available, all three could be used to represent a single object. If each representation stores an independent sample of the stimulus, and observers simply report the average of the three representations at the time of test, this will lead to an increase in the precision of the report. We call this the 'slots+averaging' model. Note that storing a single object in multiple
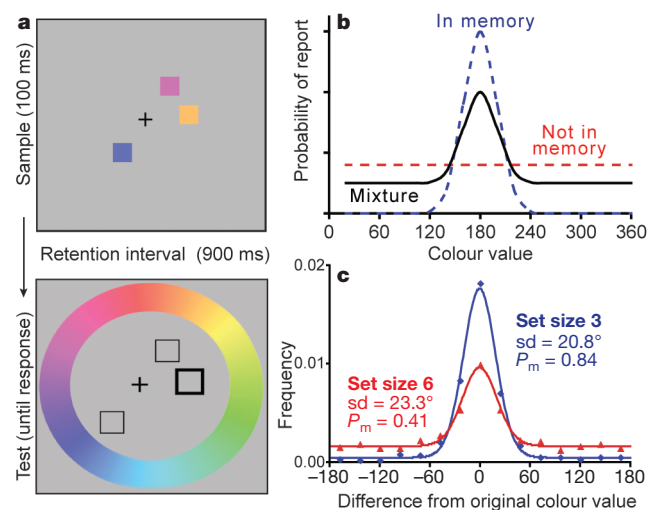
Figure 1 | Experimental approach and results of experiment 1. a, Colour recall task. **b**, Mixture model of performance, showing the probability of reporting each colour value given a sample colour at $180°$. When the probed item is present in memory, the reported colour tends to be near the original colour (blue broken line). When the probed item is not present in memory, the observer is equally likely to report any colour value (red broken line). When collapsed across trials, the data comprise a mixture of these two trial types (solid line), weighted by the probability that the probed item was stored in memory. **c**, Results of experiment 1 ($N = 8$). P_m and s.d. are defined in the text.

[1]Department of Psychology, University of Iowa, Iowa City, Iowa 52242, USA. [2]Center for Mind & Brain, University of California, Davis, California 95618, USA.

slots would be a rational strategy, and this sort of averaging is common in models of perception[8–10].

For both the slots+resources and slots+averaging models, s.d. will be improved when the set size is reduced below the number of available slots. Moreover, both models predict that this improvement will follow a square root function (see Supplementary Notes). This is exactly what was observed in experiment 2 (Fig. 2), in which observers ($N = 8$) were presented with 1, 2, 3 or 6 objects: s.d. increased as the set size increased from 1 to 3 but then remained constant as the set size increased to 6. In contrast, P_m declined very slowly as set size increased from 1 to 3 and then decreased suddenly at set size 6. This pattern of results can be explained quantitatively by both the slots+resources model (adjusted $r^2 = 0.96$) and the slots+averaging model (adjusted $r^2 = 0.99$) (see Fig. 2 and Supplementary Notes), but it differs significantly from the predictions of a pure resource model ($P < 0.001$, χ^2 test).

The slots+resources and slots+averaging models make different predictions about the range over which precision can vary. Specifically, the slots+resources model posits that the majority of resources can be devoted to one representation (leading to a very small s.d.), leaving 'only a few drops' of resources for other representations (leading to a very large s.d.). In contrast, the slots+averaging model posits that the observed s.d. is never worse than the s.d. of a single slot and is never better than the s.d. for a single slot divided by the square root of the number of slots. To distinguish between these models, experiment 3 ($N = 22$) used a line in the sample array to cue one of four coloured squares (Fig. 3a). The cued square was probed on 70% of the trials, and each uncued square was probed on 10% of trials. Neutral trials were also included, in which all four locations were cued. The cue was simultaneous with the sample array so that it would not influence perceptual processing[11], and the duration of the sample array was increased to 300 ms to provide adequate time for resource allocation[10,12].

The slots+resources model predicts that observers will devote the lion's share of resources to the cued item, leading to a large difference in s.d. between valid, neutral and invalid trials, but only a small difference in P_m. In contrast, the slots+averaging model predicts that observers will devote most of their slots to the cued location, which would lead to a large difference in P_m between valid and invalid trials. This should also lead to a somewhat smaller s.d. on valid trials than on neutral trials because of the benefits of averaging. However, this should lead to no difference in s.d. between neutral and invalid trials, because a given item receives either 0 or 1 slots on both neutral and invalid trials.

We found that P_m was substantially greater on valid trials than on invalid trials ($F(1,21) = 203.87$, $P < 0.001$; Fig. 3a), demonstrating that the observers attempted to maximize performance for the cued item by devoting more slots to it. s.d. was slightly but significantly smaller on valid trials than on neutral trials ($F(1,21) = 13.49$, $P < 0.001$), and the magnitude of this difference was within the small range that can result from averaging slots. In addition, s.d. was virtually identical on neutral and invalid trials ($F < 1$), indicating
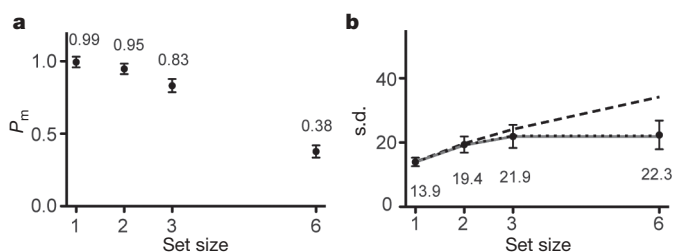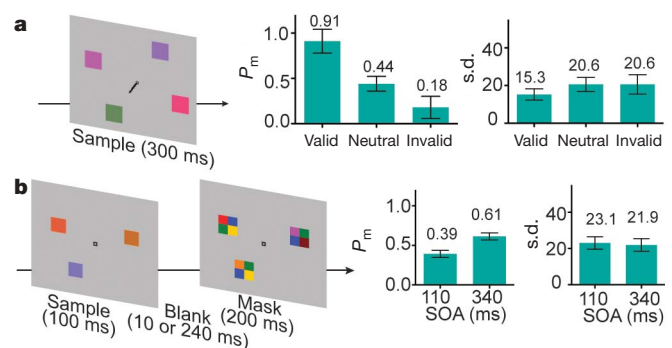
that the improvement in s.d. on valid trials was not achieved by taking resources away from the uncued items. Thus, despite the fact that the cued item was seven times more likely to be probed than each uncued item, s.d. was only slightly improved for the cued item (compared to the neutral trials) and s.d. was not reduced for the uncued items (compared with the neutral trials). It does not appear to be possible to provide a representation with 'only a few drops' of resources and thereby produce an imprecise representation.

Computational neuroscience theories suggest that an all-or-none, fixed-resolution encoding process may be required to create durable representations that can survive new sensory inputs[13,14]. To assess the encoding process, experiment 4 ($N = 8$) used a masking manipulation that emulates the masking effects of eye movements in natural vision. Specifically, we presented masks at the locations of the coloured squares either 110 or 340 ms after the onset of the squares (Fig. 3b). At these intervals, masks interfere with working memory encoding but not with perceptual processing[15]. If working memory representations gradually become more precise over time, then presenting a mask array at an early time point could potentially reveal the existence of low-precision representations. If, however, the process of creating durable memory representations that can survive new visual inputs involves an all-or-none step, as suggested by studies of the 'attentional blink' phenomenon[16], then the masks will only influence P_m. We observed that decreasing the masking interval produced a large decline in P_m ($F(1,7) = 47.70$, $P < 0.001$) but no change in s.d. ($F < 1$). Thus, the creation of working memory representations that can survive new inputs involves an all-or-none step for simple objects (although it is possible that some gradual accumulation of information occurs before this step and is available in the absence of masking).

To demonstrate that the present results can generalize to other stimulus dimensions, we repeated experiments 2 and 3 with shapes rather than colours. We used shapes defined by Fourier descriptors[17], which vary along continuous quantitative dimensions. The results were largely identical to the results obtained for colour, with approximately the same P_m for these shapes as for the simple colours in experiment 2 (see Supplementary Notes). Most notably, s.d. did not increase as the set size increased from three to six items and was virtually identical for neutral and invalid trials. Because the Fourier descriptor method provides a mathematically[17], perceptually[18,19] and neurally[20] meaningful way to describe shapes of any complexity, this analytic approach could be used to determine whether the present pattern of results would be obtained with more complex objects. Object complexity can have a large impact on performance in change detection tasks[7], but this may reflect greater sample-test similarity for complex objects[21] or the need to store each

Figure 2 | P_m and s.d. results from experiment 2 ($N = 8$). **a**, P_m; **b**, s.d.; the numbers within the panels provide the means. In **b**, the lines show the s.d. predictions of a pure resource model (black dashed line), the slots+averaging model (grey solid line), and the slots+resources model (black dotted line). Error bars show within-subjects 95% confidence intervals[26].

Figure 3 | Stimuli and results from experiments 3 and 4. **a**, Experiment 3 ($N = 22$), which included valid, neutral and invalid trials. The cue appeared simultaneously with the sample array. **b**, Experiment 4 ($N = 8$), in which a mask array followed the sample array with a stimulus onset asynchrony (SOA) of 110 or 340 ms. A colour wheel and probe array appeared at the end of the trial, 900 ms after sample offset. Error bars show within-subjects 95% confidence intervals[26], and numbers within the panels provide the means.

part of a multipart object in a separate slot[22,23]. Alternatively, complex objects may require some kind of limited resource that is not needed for the simple objects studied here.

Together, the present experiments resolve an issue that has been debated for decades[4,5,24,25], showing that a model with a small set of discrete, fixed-resolution representations can provide a quantitative account of memory performance across a broad range of experimental manipulations. This model does not completely eliminate the concept of resources, because the slots themselves are a type of resource. However, the slots+averaging model defines exactly what the resource is and describes strict limits on how flexibly this resource can be allocated.

METHODS SUMMARY

The stimuli and task are shown in Fig. 1a. Subjects viewed a sample array and then, following a brief delay, reported the colour of one item from this array (indicated by a thick outlined box) by clicking on a colour wheel. There is good agreement between this procedure and the more commonly used change-detection procedure (see Supplementary Notes).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Baddeley, A. D. *Working Memory* (Clarendon, Oxford, 1986).
2. Miyake, A. & Shah, P. (eds) *Models of Working Memory* (Cambridge Univ. Press, Cambridge, UK, 1999).
3. Luck, S. J. & Vogel, E. K. The capacity of visual working memory for features and conjunctions. *Nature* **390**, 279–281 (1997).
4. Frick, R. W. Issues of representation and limited capacity in the visuospatial sketchpad. *Br. J. Psychol.* **79**, 289–308 (1988).
5. Wilken, P. & Ma, W. J. A detection theory account of change detection. *J. Vis.* **4**, 1120–1135 (2004).
6. Prinzmetal, W., Amiri, H., Allen, K. & Edwards, T. Phenomenology of attention: I. Color, location, orientation, and spatial frequency. *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 261–282 (1998).
7. Alvarez, G. A. & Cavanagh, P. The capacity of visual short-term memory is set both by information load and by number of objects. *Psychol. Sci.* **15**, 106–111 (2004).
8. Palmer, J. Attentional limits on the perception and memory of visual information. *J. Exp. Psychol. Hum. Percept. Perform.* **16**, 332–350 (1990).
9. Newsome, W. T. *et al.* in *The Cognitive Neurosciences* (ed. Gazzaniga, M. S.) 401–414 (MIT Press, Cambridge, Massachusetts, 1995).
10. Bonnel, A. M. & Miller, J. Attentional effects on concurrent psychophysical discriminations: Investigations of a sample-size model. *Percept. Psychophys.* **55**, 162–179 (1994).
11. Vogel, E. K., Woodman, G. F. & Luck, S. J. Pushing around the locus of selection: Evidence for the flexible-selection hypothesis. *J. Cogn. Neurosci.* **17**, 1907–1922 (2005).
12. Luck, S. J., Hillyard, S. A., Mouloua, M. & Hawkins, H. L. Mechanisms of visual-spatial attention: Resource allocation or uncertainty reduction? *J. Exp. Psychol. Hum. Percept. Perform.* **22**, 725–737 (1996).
13. Wang, X.-J. Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* **24**, 455–463 (2001).
14. Raffone, A. & Wolters, G. A cortical mechanism for binding in visual working memory. *J. Cogn. Neurosci.* **13**, 766–785 (2001).
15. Vogel, E. K., Woodman, G. F. & Luck, S. J. The time course of consolidation in visual working memory. *J. Exp. Psychol. Hum. Percept. Perform.* **32**, 1436–1451 (2006).
16. Sergent, C. & Dehaene, S. Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychol. Sci.* **15**, 720–728 (2004).
17. Zahn, C. T. & Roskies, R. Z. Fourier descriptors for plane closed curves. *IEEE Trans. Comput. C* **21**, 269–281 (1972).
18. Alter, I. & Schwartz, E. L. Psychophysical studies of shape with Fourier descriptor stimuli. *Perception* **17**, 191–202 (1988).
19. Cortese, J. M. & Dyre, B. P. Perceptual similarity of shapes generated from Fourier descriptors. *J. Exp. Psychol. Hum. Percept. Perform.* **22**, 133–143 (1996).
20. Schwartz, E. L., Desimone, R., Albright, T. D. & Gross, C. G. Shape recognition and inferior temporal neurons. *Proc. Natl Acad. Sci. USA* **80**, 5776–5778 (1983).
21. Awh, E., Barton, B. & Vogel, E. K. Visual working memory represents a fixed number of items regardless of complexity. *Psychol. Sci.* **18**, 622–628 (2007).
22. Xu, Y. Limitations of object-based feature encoding in visual short-term memory. *J. Exp. Psychol. Hum. Percept. Perform.* **28**, 458–468 (2002).
23. Sakai, K. & Inui, T. A feature-segmentation model of short-term visual memory. *Perception* **31**, 579–590 (2002).
24. Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* **63**, 81–97 (1956).
25. Simon, H. A. How big is a chunk? *Science* **183**, 482–488 (1974).
26. Cousineau, D. Confidence intervals in within-subjects designs: A simpler solution to Loftus and Masson's method. *Tutor. Quant. Methods Psychol.* **1**, 42–45 (2007).
27. Myung, I. J. Tutorial on maximum likelihood estimation. *J. Math. Psychol.* **47**, 90–100 (2003).
28. Fisher, N. I. *Statistical Analysis of Circular Data* (Cambridge Univ. Press, New York, 1993).

METHODS

Subjects. Eight subjects between 18 and 35 years old participated in each colour memory experiment except the cuing experiment, in which 22 subjects participated owing to the low probability of the invalid trials. Subjects provided informed consent and received course credit or monetary compensation. All reported having normal colour vision and normal or corrected-to-normal visual acuity.

Experiments. The stimuli in all experiments were presented on a CRT monitor with a grey background ($15.5\,\mathrm{cd\,m^{-2}}$) at a viewing distance of 57 cm. The monitor was calibrated with a Tektronix J17 LumaColour colorimeter. Each coloured square in the sample array subtended $2 \times 2°$ of visual angle. Each square was centred on an invisible circle with a radius of $4.5°$. The positions were randomly chosen from a set of eight locations equally spaced along the circle. The colour wheel was $2.2°$ thick and was centred on the monitor with a radius of $8.2°$. It consisted of 180 colour values that were evenly distributed along a circle in the CIE $L*a*b*$ colour space. This circle was centred in the colour space at ($L = 70$, $a = 20$, $b = 38$) with a radius of 60. Its centre was chosen to maximize its radius and therefore the discriminability of the colours. All colours had equal luminance and varied mainly in hue and slightly in saturation. The sample array colours were randomly selected from this set of colours. The colour wheel was presented at a random rotation on each trial to minimize contributions from spatial memory.

In the basic paradigm (experiments 1 and 2), each trial consisted of a 100-ms sample array followed by a 900-ms blank delay period and then a probe display that remained present until a response was made (Fig. 1a). The probe display contained the colour wheel and an outlined square at the location of each item from the sample array. One of these squares was thicker ($0.20°$) than the others ($0.04°$), which cued the subject to recall the colour of the corresponding item from the sample array by clicking the appropriate colour on the colour wheel with the computer mouse. Accuracy was stressed, and the responses were not timed. Except as noted below, 150 trials were tested in each experimental condition (for example, each set size). The different trial types (for example, different set sizes) were presented in an unpredictable order in each experiment.

In the control experiment that involved varying the level of perceptual noise (Supplementary Fig. 1), the set size was held constant at three items and the duration of the sample array was reduced to 30 ms to ensure that the masks would be effective. Each coloured square in the sample array was covered with a set of either 75 or 150 simultaneously presented coloured dots, randomly distributed over a circular region with a diameter of $4.4°$ that was centred on the coloured square. Each dot subtended $0.2 \times 0.2°$ of visual angle and was drawn in a colour that was randomly sampled from the set of 180 colour values used for the coloured squares.

In the cuing experiment (Fig. 3a), the sample display contained a $1.6°$ cue line extending from fixation towards one of the four coloured squares (on valid and invalid trials) or four lines extending towards all four squares (on neutral trials). The duration of the sample display was increased to 300 ms in this experiment to provide the observers sufficient time to shift attention to the cued item; the interval between sample onset and probe onset remained 900 ms. Each observer received 350 valid trials, 150 invalid trials (50 per uncued location) and 150 neutral trials. These trial types were randomly intermixed.

In the backward masking experiment (Fig. 3b), the sample array always contained three items. An array of masks was presented 110 or 340 ms after the onset of the sample array, with a duration of 200 ms. The interval between sample offset and probe onset remained constant at 900 ms. Each mask consisted of a 2×2 arrangement of coloured squares, each of which measured $0.55 \times 0.55°$. Each mask was centred at the location of one of the three items in the sample display.

Data analysis. The data from a given observer in the colour experiments consisted of a set of distances between the reported colour value and the original colour value in each condition, which reflects the degree of error in the reported colour. Histograms of these error values were used to visualize the distribution of responses (as in Fig. 1c). Maximum likelihood estimation[27] was used to decompose the data from each subject in each condition into three parameters that represent a mixture of a uniform distribution of errors (for trials on which the probed item was not encoded in memory) and a von Mises distribution of errors (for trials on which the probed item was encoded). The von Mises distribution is the circular analogue of the gaussian distribution and was used because the tested colour space was circular[28]. The uniform distribution was represented by a single parameter, P_m, which is the probability that the probed item was present in memory at the time of the probe (which is inversely related to the height of the uniform distribution). The von Mises distribution was represented by two parameters, its mean (μ) and its standard deviation (s.d.). μ reflects any systematic shift of the distribution away from the original colour value. No systematic shifts were expected or observed in any of the present experiments, so this parameter will not be considered further. s.d. reflects the width of the distribution of errors on trials when the probed item was encoded in memory, which in turn reflects the precision or resolution of the memory representation.

The slots+averaging model was fitted to the estimated P_m and s.d. parameters in the experiment in which set sizes 1, 2, 3 and 6 were tested. We computed the total number of slots (K_i) by multiplying P_m by the set size (using the data from set size 3). We then assumed that the slots were randomly distributed among the available items in the sample array, allowing multiple slots to be assigned to a given object if the set size was lower than the number of slots. The s.d. at set size 3 was used to estimate the precision of a single slot. The s.d. from a set of N samples is equal to the s.d. from a single sample divided by the square root of N (see Supplementary Notes). Thus, by knowing the s.d. of a single slot and the average number of slots assigned to the probed item in a given condition, it is possible to predict the s.d. for that condition.

In the slots+resources model, the s.d. at set size 1 was used to estimate the maximum precision when all resources are devoted to a single object. For modelling the data from larger set sizes, the s.d. simply increases as a function of the square root of the number of objects being represented, up to the number of slots (which is estimated as in the slots+averaging model). For simple manipulations of set size, the predictions of the slots+resources model are equivalent to those of the slots+averaging model except that the s.d. values are estimated on the basis of the data at set size 1 rather than the data at set size 3.