Brief Communication

Hippocampus contributes to the maintenance but not the quality of visual information over time

David E. Warren,[1] Melissa C. Duff,[1,2] Neal J. Cohen,[4,5,6] and Daniel Tranel[1,3]

[1]*Department of Neurology, Carver College of Medicine,* [2]*Department of Communications Sciences and Disorders,* [3]*Department of Psychology, University of Iowa, Iowa City, Iowa 52242, USA;* [4]*Department of Psychology,* [5]*Beckman Institute,* [6]*Neuroscience Program, University of Illinois at Urbana-Champaign, Champaign Illinois 61801, USA*

The hippocampus has recently been implicated in the brief representation of visual information, but its specific role is not well understood. We investigated this role using a paradigm that distinguishes quantity and quality of visual memory as described in a previous study. We found that amnesic patients with bilateral hippocampal damage ($N = 5$) were less likely to remember test stimuli than comparison participants despite a brief maintenance interval (900 msec). However, estimates of memory quality were similar for all groups. Our findings suggest that the hippocampus contributes to brief maintenance of visual information but does not contribute to the quality of that information.

[Supplemental material is available for this article.]

The ability to retain and integrate visual information across short intervals is critical to everyday life, demonstrated whenever one examines a painting, reads a manuscript, or remembers that a stoplight is still red. The psychological processes and neural mechanisms of these briefly held visual representations have been investigated (Jonides et al. 2008; Chun et al. 2011), but the specific contributions of many brain regions to the maintenance of visual information are unclear. For example, while the medial temporal lobe (MTL) and hippocampus are often associated with the formation of lasting declarative memories (Scoville and Milner 1957; Cohen and Squire 1980), MTL and hippocampus have also been implicated in the representation of visual information over very short intervals in animal models (Eacott et al. 1994; Murray and Bussey 1999; Bussey et al. 2002; Cowell et al. 2006) and more recently in humans (Olson et al. 2006; Barense et al. 2007; Lee et al. 2012; Warren et al. 2012). However, it is not clear whether the hippocampus and MTL contribute to the maintenance of visual representations, the quality of visual representations, or both.

Neuropsychological studies have shown that damage to MTL structures including the hippocampus causes impairment in performance on many tasks at short delays (Barense et al. 2005, 2007; Lee et al. 2005a,b; Hannula et al. 2006; Lee and Rudebeck 2010; Warren et al. 2010, 2011, 2012; Kurczek et al. 2013; Watson et al. 2013). Typically, these investigations have used binary yes/no or forced choice recognition tasks that cannot address how hippocampal damage might change the quality of mental representations. Tasks that permit graded responses have shown that MTL and hippocampal damage increase the variance of responses over time relative to neurologically normal participants (Sidman et al. 1968; Warren et al. 2010), but have not explained any underlying representational changes.

Zhang and Luck (2008) developed a theory and method sufficient to inform this issue by beginning with the premise that mental representations are inherently noisy. From this perspective, representational changes in short-delay tasks (whether in healthy participants or those with hippocampal damage) could be due to: added noise in mental representations that reduce their quality, leading to test–time mismatch with the original stimulus; increased probability that a stimulus is completely forgotten; or a combination of these phenomena (Fig. 1A). Studies of visual working memory in healthy participants suggest that visual representations follow the second course, disappearing from memory rather than decreasing in quality over time (Zhang and Luck 2009, but see Bays et al. 2009). Critically, the method of Zhang and Luck (2008, 2009, 2011) supports independent estimation of the probability of a tested item being represented in memory and the quality of memory representations, providing significant advantages over binary response tasks.

Here, we evaluated the necessity of hippocampus for the brief maintenance of simple visual information using a neuropsychological approach. We tested neurological patients with amnesia ($N = 5$, 1F/4M); patients with brain damage excluding MTL and hippocampus ("BDC"; $N = 14$, 6F/8M, 1 M later excluded for color blindness); and healthy comparison participants ("NC"; $N = 19$, 9F/10M) of similar age and educational attainment (Table 1; $P > 0.05$) using a task that provides insight into the quantity and quality of visual memory representations at short delays (Fig. 1B; Zhang and Luck 2008). Participants saw 1, 3, or 6 color stimuli presented briefly (100 msec) and 900 msec later responded to a memory probe in a particular location by selecting the color previously occupying that location from a color ring. Based on prior research (Sidman et al. 1968; Downes et al. 1998; Warren et al. 2010), we hypothesized that amnesic patients would show a broader distribution of responses than comparisons, reflecting reduced retention of visual information versus comparisons.

Corresponding author: david-e-warren@uiowa.edu
Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.037127.114.

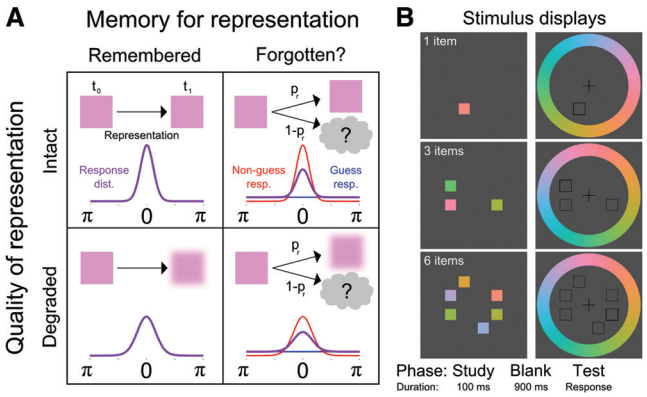

Figure 1. Model assumptions and trial sequence. (*A*) Based on the model of Zhang and Luck (2008), we hypothesized that briefly maintained visual representations could change in two dissociable ways. Representations could be remembered or forgotten (*left* and *right* columns), and the quality of representations could remain intact or be reduced (*top* and *bottom* rows). Each panel diagrams a combination of forgetting and degradation of the visual representation (*top*), and the associated distribution of responses (*bottom*) around the target value (0, with maximum response error of π). Expected response distributions (purple) could change in two ways depending on changes to underlying representations: reduced quality would yield a broader distribution of responses; while forgetting some representations entirely would yield a hybrid of a uniform distribution reflecting guesses (blue) and a target-centered distribution of memory-guided responses (red). We predicted that hippocampal damage would reduce probability of memory for studied items but not degrade representations (upper *right*). (*B*) In each trial, participants saw 1, 3, or 6 color squares for 100 msec (white text was not presented). After a 900-msec blank interval, the target location was indicated with a thick open square. Participants selected the color that was seen in that location from the color wheel.

Importantly, the sensitive nature of our task and analysis also allowed us to form novel hypotheses regarding the specific nature of the impairment. Following the well-established role of the MTL and hippocampus in memory, we predicted that the visual representations of amnesic patients would be more susceptible to forgetting, but that representations which were available at test would be distributed similarly to those of comparisons.

Task materials were based on those described by Zhang and Luck (2008). Displays consisted of closed squares of specific colors that subtended 2° of visual angle (horizontally and vertically), open black squares subtending 2.04° and 2.2° of visual angle, and a colored ring with an inner radius of 7.1° of visual angle and an outer radius of 9.3° of visual angle (Fig. 1B). For the colored ring, 180 equal-luminance colors were selected from CIE $L^*a^*b^*$ color space by sampling the space around $L = 70$, $a = 0$, $b = 13$ in 180 even steps around the circumference of a circle in the a^*b^* plane with radius 45. All values were converted to RGB and checked for compatibility with that color space. Colors for the closed square stimuli were drawn from this 180-color spectrum. Visual stimuli were presented at a distance of 50 cm on a 21-in LCD monitor with a vertical refresh rate of 60 Hz (MultiSync 2190UXi, NEC Corporation of America). Responses were made with a computer mouse.

Our procedure adapted the Zhang and Luck paradigm (2008). Participants were seated in front of a computer display. At the beginning of each block, written instructions were presented on the screen: "Colored squares will briefly appear near the center of the screen. Remember all of the colors. When one position is cued, indicate the color from that position." Complementary verbal instructions by the experimenter emphasized key task components, and participant comprehension was evaluated. A practice block (15 three-item trials) preceded the main test phase. The main test phase consisted of three blocks containing 150 trials including 50 trials each for 1, 3, and 6 items in a unique random order.

Participants initiated each trial with a mouse click; a central fixation cross changed color, and ~1 sec later the trial began (the stimulus onset asynchrony jittered by ±125 msec). The trial sequence (Fig. 1B) was: a study display containing one or more color squares (100 msec); a blank display (900 msec); and the test display (presented until response) which included a mouse cursor, open squares surrounding each position that previously held a colored square, a cue in the form of open square with a thicker outline surrounding the test location, and the color ring. Participants indicated which color had been presented in the cued location by clicking that color on the color ring, and guessed if unsure.

Based on the distribution of test–time responses, three parameters were estimated for each participant: p_r, the probability that the probe item was in memory at test (i.e., the "quantity" of information in memory); k, the concentration of the response distribution around μ (i.e., the "quality" of information in memory); and μ, the mean of the response distribution. Parameter estimation was based on previously reported methods (Zhang and Luck 2008), and is described in the Supplemental Material. All trials with response times ≤15 sec were used in the parameter estimation procedure. Group differences were evaluated for the three parameters (p_r, κ, μ). No significant differences were found for μ (see Supplemental Material). In order to address the possibility that nontarget items significantly influenced response distributions, we also estimated parameters for an alternative model (Bays et al. 2009). Results were generally consistent with the main findings; the alternative approach and results are described in the Supplemental Material. Between-group differences for all dependent variables were tested using repeated-measures ANOVA implemented as a hierarchical linear model with participants as a random effect, group as a between-participants fixed effect, and number of items as a within-participants fixed effect. Planned between-group and between-condition comparisons were conducted using nonpaired, equal-variance *t*-tests. Permutation tests of the planned comparisons are reported as p_{perm} and were calculated as follows: bootstrapped distributions were created by assigning group membership to the data in 10^5 randomly selected permutations, recording the statistic value for each permutation, and determining the percentile rank of the observed statistic value in the bootstrapped distribution. Effect size was measured with an unbiased variant of Cohen's d that accounts for small sample sizes (d_{unb}) (Grissom and Kim 2012, p. 70). Response time was not a dependent variable of interest, but a similar, exploratory analysis is presented in the Supplemental Material and Table S2.

The probability of an item being present in memory at test (p_r) differed between groups and was significantly reduced in amnesic patients (Fig. 2A; Supplemental Table S1). Between-group differences in p_r were significant [$F_{(2,34)} = 4.597$, $P < 0.001$], and planned comparisons between groups for each number of items showed that amnesic patients were significantly impaired relative to the NC and BDC groups in the three-item condition [NC, $T_{(22)} = 2.879$, $P = 0.009$, $p_{\text{perm}} = 0.006$; BDC, $T_{(16)} = 3.020$, $P = 0.008$, $p_{\text{perm}} = 0.005$], at least marginally impaired in the six-item condition [NC, $T_{(22)} = 2.070$, $P = 0.050$, $p_{\text{perm}} = 0.015$; BDC, $T_{(16)} = 2.152$, $P = 0.047$, $p_{\text{perm}} = 0.029$], but no more than marginally impaired in the one-item condition [NC, $T_{(18)} = 1.878$, $P = 0.077$, $p_{\text{perm}} = 0.073$; BDC, $T_{(16)} = 1.611$, $P = 0.127$, $p_{\text{perm}} = 0.133$]. The lack of a significant difference for one item may have reflected a ceiling effect as all groups had values of p_r near 1 (see Supplemental Table S1 and Discussion). The NC and BDC groups did not differ in any condition [one item,

Table 1. Demographic and neuropsychological data characterizing participants

ID	Age	Sex	Edu.	Eti.	Chr.	Hand	FSIQ	VIQ	PIQ	DS	GMI	AVLT	CFT C/R	BVRT	HcV
1846	49	F	14	An./SE	19	100	84	89	79	10	57	7/3	28/6	5	−4.23*
1951	60	M	16	HSE	32	100	106	105	106	9	57	9/1	32/4	6	≫
2308	56	M	16	HSE	13	−100	98	95	92	9	45	5/0	32/0	6	≫
2363	56	M	18	An.	14	100	98	112	83	8	73	8/0	26/5	6	−2.64*
2563	58	M	16	An.	13	−80	94	91	98	14	63	10/4	36/7	7	NA
Amn ($N = 5$)	55.8 (4.1)	4M 1F	16.0 (1.4)	–	18.2 (8.1)	–	96.0 (8.0)	98.4 (9.8)	91.6 (11.0)	10.0 (2.3)	59.0 (10.2)	7.8/1.6 (1.9/1.8)	30.8/3.9 (4.4/2.7)	6.0 (0.7)	–
BDC ($N = 13$)	61.5 (8.9)	7M 6F	14.1 (2.0)	–	11.1 (8.3)	–	110.7 (9.8)	108.8 (11.2)	110.5 (11.1)	10.5 (2.5)	–	12.8/10.2 (2.3/3.2)	31.6/18.0 (3.1/7.0)	7.9 (2.7)	–
NC ($N = 19$)	53.1 (6.7)	10M 9F	15.7 (1.9)	–	–	–	–	–	–	–	–	–	–	–	–

Individual scores are presented for each participating amnesic patient, followed by amnesic (Amn) group means, brain-damaged comparison (BDC) group means, and healthy normal comparison (NC) group means (standard deviations in parentheses). The significant memory impairment of the amnesic group is evident in several neuropsychological measures. Note that these scores may reflect updated test results based on periodic case reviews, and are contemporaneous with this study. See Lezak et al. (2012) for further information about individual measures. Abbreviations: Age, years; Edu., education, years; Chr., chronicity, years since injury; Hand, handedness (+100 = fully right handed, −100 = fully left handed); Eti., etiology; Anoxia/An., anoxic/ischemic episode, SE, status epilepticus, HSE, herpes simplex encephalitis; FSIQ, WAIS-III full-scale IQ (WAIS-IV was used for some BDC patients); VIQ, verbal IQ; PIQ, performance IQ; DS, WAIS 3/4 Digit Span; WMS-III GMI, general memory index; AVLT, Rey Auditory Verbal Learning Task, trial 5/30-min delay; CFT, complex figure task copy/recall; BVRT, Benton visual retention test number correct; HcV, bilateral hippocampal volumes per Allen et al. (2006). Volumes are expressed in Studentized residuals relative to normative expectations: *, reported by Allen et al. (2006); ≫, residual value not available, but near-complete bilateral hippocampal lesion in the context of larger brain lesions (see Feinstein et al. 2010; Cavaco et al. 2012); NA, volumetric measurements unavailable due to contraindications for MRI (e.g., pacemaker).

$T_{(26)} = 0.445$, $P = 0.660$, $p_{perm} = 0.328$; three items, $T_{(30)} = 1.047$, $P = 0.304$, $p_{perm} = 0.150$; six items, $T_{(30)} = 1.089$, $P = 0.285$, $p_{perm} = 0.143$].

In addition to between-group effects, p_r was affected by the number of items presented [$F_{(2,64)} = 363.874$, $P < 0.001$], but the interaction of the group and number-of-items factors was not significant [$F_{(4,64)} = 1.428$, $P = 0.235$]. For all groups the pattern was similar (Fig. 2A): p_r was greatest for one item; relatively less for three items; and least for six items. Planned within-group comparisons between p_r for presentations of one versus three items and three versus six items showed that this pattern was significant for all groups [one item versus three items, each $T > 4.5$, each $P < 0.0025$, each $p_{perm} < 0.005$, each $d_{unb} > 1.9$; three items versus six items, each $T > 4.0$, each $P < 0.005$, each $p_{perm} < 0.001$, each $d_{unb} > 2.4$]. Thus, when more stimuli were presented, the probability of any given stimulus being present in memory at test was reduced.

The quality of memory representations (κ) was influenced by the number of items in a display, but did not differ between groups (Fig. 2B; Supplemental Table S1). The NC, BDC, and amnesic groups all had similar κ [$F_{(2,34)} = 0.229$, $P = 0.796$], and planned comparisons between groups for each number of items found no significant differences [each $T < 1.4$, each $P > 0.19$, each $p_{perm} > 0.09$]. Meanwhile, the number of items presented had a significant effect on κ [$F_{(2,64)} = 17.478$, $P < 0.001$]. Planned comparisons revealed the same pattern for each group: displays of one item produced the greatest value of κ [one versus three items, each $T > 3.2$, each $P < 0.025$, each $p_{perm} < 0.005$]; while presentations of three and six items produced values of κ that were lower and not statistically different from each other [three versus six items, each $T < 1.7$, each $P > 0.10$, each $p_{perm} > 0.501$]. There was no significant interaction of the group by condition [$F_{(4,64)} = 0.427$, $P > 0.789$]. This pattern of higher quality representations for one item than for either of the larger sets of items may reflect previously reported characteristics of visual working memory (Zhang and Luck 2008).

The observed deficit in brief maintenance of visual information by amnesic patients with hippocampal damage could be attributed to reduced memory capacity, reduced ability to maintain information over time, a combination of these factors, or still further causes. Our findings are compatible with a previously hypothesized role for the hippocampus in the on-line processing of visual information (Gallegos et al. 2006; Barense et al. 2007; Warren et al. 2011, 2012), and congruent with suggestions that short-term and long-term memory systems may not be neurally dissociable (Ranganath and Blumenfeld 2005). Moreover, we suggest that the role of the hippocampus in visual representation is inherently mnemonic and relational. For example, relational memory theory (Cohen and Eichenbaum 1993; Eichenbaum and Cohen 2001, Moses and Ryan 2006, Ranganath 2010; Eichenbaum and Cohen 2014) predicts that the hippocampus is necessary for the binding of arbitrarily related information (e.g., color and spatial location) irrespective of timescale. What our current methodology identifies as outright loss of information may include some responses in the three- and six-item conditions that

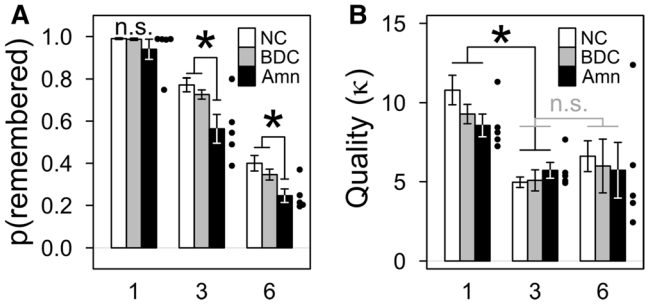

Figure 2. Damage to the hippocampus and MTL reduced the probability that an item would be remembered without altering the quality of memory representations. (*A*) Group means for the probability of remembering the tested item. Amnesic patients (Amn) were significantly (*) less likely to remember items at test overall and specifically for the three- and six-item conditions versus both comparison groups. Error bars show SEM, and the performance of individual amnesic patients is indicated by points. In the three- and six-item conditions patient 1846 performed better than the other amnesic patients and near the comparison means; more information and detailed parameter fits are provided in Supplemental Table S3. (*B*) Group means for the quality of remembered representations (i.e., discounting forgetting) were similar for all item conditions, but quality was significantly (*) greater in the one-item condition than the three- and six-item conditions. Error bars and points as in panel *A*.

involve color–location association errors on the part of participants, and relational memory theory predicts that hippocampal damage would increase these errors (Watson et al. 2013). In this context, the relatively preserved performance of amnesic patients in the one-item condition could be due to a lack of relational demands. However, a supplemental analysis using the model of Bays et al. (2009), which attempts to account for responses driven by nontarget items, did not provide strong evidence for increased color–location association errors by amnesic patients (see Supplemental Results).

We suggest that the observed impairment in the brief maintenance of visual information by amnesic patients is due to impairments in on-line processing of relations due to hippocampal damage, but other data and interpretations are relevant. Zhang and Yonelinas (2012) tested a mixed group of unilateral temporal lobectomy and anoxic amnesic patients using similar methodology, and reported a change in the quality of memory representations (i.e., decreased κ) rather than the probability of memory (decreased p_r). Both studies clearly show that hippocampal damage can impair performance on the Zhang and Luck (2008) task; different patterns of results could be attributable to patient anatomy or details of task implementations. Meanwhile, Jeneson et al. (2010, 2012) and Jeneson and Squire (2011) have suggested that deficits at similar timescales are attributable to the inability of amnesic patients to remember information exceeding the capacity of short-term memory because they lack normal declarative memory systems. We note that our task used a brief maintenance interval (900 msec) and that the amnesic group showed an impairment that was significant for dislays containing as few as three items. Jeneson et al. (2012) have previously described displays with these characteristics as within the capacity of short-term memory, and we concur with that description. We attribute our finding of impairment in this context to the sensitive nature of our experimental methodology.

Despite our robust findings, the study had some limitations. As in most neuropsychological investigations studying severely amnesic patients, our sample size was relatively small. However, the study had enough power to uncover significant differences, and our main findings had substantial effect sizes. Interestingly, while we observed impairment for the amnesic group that was greatest for presentations of three and six items, presentations of one item did not reliably produce impairment. Single items may have been maintained normally by the amnesic group, but it is possible that differences in the maintenance of a single item were obscured by a ceiling effect (especially among comparison participants). Further exploration of the parameter space in this task could attempt to address ceiling effects by including a two-item condition or a reducing item exposure time.

In summary, we found that hippocampal damage was related to reduced probability of remembering briefly maintained mental representations of visual information, indicating that the hippocampus normally makes important contributions to remembering visual information over very short intervals. Our results suggest that future investigations of visual representations in hippocampal amnesic patients could benefit from using graded rather than binary response designs in order to collect rich response distributions. We predict that populations with damage or dysfunction of hippocampus will show deficits at short delays when tested with stimuli of visual or other modalities, further demonstrating the contributions of hippocampus to brief representation and on-line processing.

Acknowledgments

We thank the following funding agencies: NINDS P01 NS19632 (DT); NIDCD R01 DC011755 (M.C.D.); NIMH R01 MH062500 (N.J.C. and D.E.W.). We thank the participating patients and their families for facilitating this investigation. We thank Samuel H. Jones and Kendra Schmitt for their assistance with this project.

Competing interest statement: The authors report no perceived or actual conflicts of interest.

References

Allen JS, Tranel D, Bruss J, Damasio H. 2006. Correlations between regional brain volumes and memory performance in anoxia. *J Clin Exp Neuropsychol* **28:** 457–476.

Barense MD, Bussey TJ, Lee ACH, Rogers TT, Davies RR, Saksida LM, Murray EA, Graham KS. 2005. Functional specialization in the human medial temporal lobe. *J Neurosci* **25:** 10239–10246.

Barense MD, Gaffan D, Graham KS. 2007. The human medial temporal lobe processes online representations of complex objects. *Neuropsychologia* **45:** 2963–2974.

Bays PM, Catalao RF, Husain M. 2009. The precision of visual working memory is set by allocation of a shared resource. *J Vis* **9:** 7.1–7.11.

Bussey TJ, Saksida LM, Murray EA. 2002. Perirhinal cortex resolves feature ambiguity in complex visual discriminations. *Eur J Neurosci* **15:** 365–374.

Cavaco S, Feinstein JS, van Twillert H, Tranel D. 2012. Musical memory in a patient with severe anterograde amnesia. *J Clin Exp Neuropsychol* **34:** 1089–1100.

Chun MM, Golomb J, Turk-Browne NB. 2011. A taxonomy of external and internal attention. *Annu Rev Psychol* **62:** 73–101.

Cohen NJ, Eichenbaum H, ed. 1993. Memory, amnesia, and the hippocampal system. The MIT Press, Cambridge, MA.

Cohen NJ, Squire LR. 1980. Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science* **210:** 207–210.

Cowell RA, Bussey TJ, Saksida LM. 2006. Why does brain damage impair memory? A connectionist model of object recognition memory in perirhinal cortex. *J Neurosci* **26:** 12186–12197.

Downes JJ, Holdstock JS, Symons V, Mayes AR. 1998. Do amnesics forget colours pathologically fast? *Cortex* **34:** 337–355.

Eacott MJ, Gaffan D, Murray EA. 1994. Preserved recognition memory for small sets, and impaired stimulus identification for large sets, following rhinal cortex ablations in monkeys. *Eur J Neurosci* **6:** 1466–1478.

Eichenbaum H, Cohen NJ. 2001. The hippocampal memory system. In *From conditioning to conscious recollection: memory systems of the brain* (Eichenbaum H, Cohen NJ), pp. 305–343. Oxford University Press, New York, NY.

Eichenbaum H, Cohen NJ. 2014. Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron* **83:** 764–770.

Feinstein JS, Rudrauf D, Khalsa SS, Cassell MD, Bruss J, Grabowski TJ, Tranel D. 2010. Bilateral limbic system destruction in man. *J Clin Exp Neuropsychol* **32:** 88–106.

Gallegos D, Tranel D, Luck SJ. 2006. Visual working memory in patients with bilateral hippocampal damage. *13th Annual Meeting of the Cognitive Neuroscience Society*, p. 237.

Grissom RJ, Kim JJ ed. 2012. *Effect sizes for research: univariate and multivariate applications*. Routledge, New York, NY.

Hannula DE, Tranel D, Cohen NJ. 2006. The long and the short of it: relational memory impairments in amnesia, even at short lags. *J Neurosci* **26:** 8352–8359.

Jeneson A, Squire LR. 2011. Working memory, long-term memory, and medial temporal lobe function. *Learn Mem* **19:** 15–25.

Jeneson A, Mauldin KN, Squire LR. 2010. Intact working memory for relational information after medial temporal lobe damage. *J Neurosci* **30:** 13624–13629.

Jeneson A, Wixted JT, Hopkins RO, Squire LR. 2012. Visual working memory capacity and the medial temporal lobe. *J Neurosci* **32:** 3584–3589.

Jonides J, Lewis RL, Nee DE, Lustig CA, Berman MG, Moore KS. 2008. The mind and brain of short-term memory. *Annu Rev Psychol* **59:** 193–224.

Kurczek J, Brown-Schmidt S, Duff M. 2013. Hippocampal contributions to language: evidence of referential processing deficits in amnesia. *J Exp Psychol Gen* **142:** 1346–1354.

Lee ACH, Rudebeck SR. 2010. Human medial temporal lobe damage can disrupt the perception of single objects. *J Neurosci* **30:** 6588–6594.

Lee ACH, Buckley MJ, Pegman SJ, Spiers H, Scahill VL, Gaffan D, Bussey TJ, Davies RR, Kapur N, Hodges JR, et al. 2005a. Specialization in the medial temporal lobe for processing of objects and scenes. *Hippocampus* **15:** 782–797.

Lee ACH, Bussey TJ, Murray EA, Saksida LM, Epstein RA, Kapur N, Hodges JR, Graham KS. 2005b. Perceptual deficits in amnesia:

challenging the medial temporal lobe 'mnemonic' view. *Neuropsychologia* **43:** 1–11.
Lee ACH, Yeung LK, Barense MD. 2012. The hippocampus and visual perception. *Front Hum Neurosci* **6:** 91.
Lezak MD, Howieson DB, Bigler ED, Tranel D, ed. 2012. *Neuropsychological assessment*. Oxford University Press, New York.
Moses SN, Ryan JD. 2006. A comparison and evaluation of the predictions of relational and conjunctive accounts of hippocampal function. *Hippocampus* **16:** 43–65.
Murray EA, Bussey TJ. 1999. Perceptual-mnemonic functions of the perirhinal cortex. *Trends Cogn Sci* **3:** 142–151.
Olson IR, Moore KS, Stark M, Chatterjee A. 2006. Visual working memory is impaired when the medial temporal lobe is damaged. *J Cogn Neurosci* **18:** 1087–1097.
Ranganath C. 2010. A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus* **20:** 1263–1290.
Ranganath C, Blumenfeld RS. 2005. Doubts about double dissociations between short- and long-term memory. *Trends Cogn Sci* **9:** 374–380.
Scoville WB, Milner B. 1957. Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* **20:** 11–21.
Sidman M, Stoddard LT, Mohr JP. 1968. Some additional quantitative observations of immediate memory in a patient with bilateral hippocampal lesions. *Neuropsychologia* **6:** 245–254.
Warren DE, Duff MC, Tranel D, Cohen NJ. 2010. Medial temporal lobe damage impairs representation of simple stimuli. *Front Hum Neurosci* **4:** 35.
Warren DE, Duff MC, Tranel D, Cohen NJ. 2011. Observing degradation of visual representations over short intervals when medial temporal lobe is damaged. *J Cogn Neurosci* **23:** 3862–3873.
Warren DE, Duff MC, Jensen U, Tranel D, Cohen NJ. 2012. Hiding in plain view: lesions of the medial temporal lobe impair online representation. *Hippocampus* **22:** 1577–1588.
Watson PD, Voss JL, Warren DE, Tranel D, Cohen NJ. 2013. Spatial reconstruction by patients with hippocampal damage is dominated by relational memory errors. *Hippocampus* **23:** 570–580.
Zhang W, Luck SJ. 2008. Discrete fixed-resolution representations in visual working memory. *Nature* **453:** 233–235.
Zhang W, Luck SJ. 2009. Sudden death and gradual decay in visual working memory. *Psychol Sci* **20:** 423–428.
Zhang W, Luck SJ. 2011. The number and quality of representations in working memory. *Psychol Sci* **22:** 1434–1441.
Zhang W, Yonelinas AP. 2012. The influence of medial temporal lobe damage on capacity and precision in visual working memory. In Paper presented at the Annual Meeting of the the Cognitive Neuroscience Society, Chicago, IL.

Received September 26, 2014; accepted in revised form October 14, 2014.

Hippocampus contributes to the maintenance but not the quality of visual information over time

David E. Warren, Melissa C. Duff, Neal J. Cohen, et al.

Learn. Mem. 2015 22: 6-10
Access the most recent version at doi:10.1101/lm.037127.114

Supplemental Methods

Participants

All patients were drawn from the Iowa Registry of Neurological Patients ("the Registry") which contains data for patients with focal, stable brain lesions who have undergone comprehensive neuropsychological testing in the chronic (>3 months after symptom onset) epoch. All participants granted informed consent before participating and were paid for their participation. Consent procedures and task administration were conducted in accordance with the Declaration of Helsinki.

The etiology of patients 1846 and 2363 was anoxia (combined with status epilepticus in 1846), and both showed significant atrophy of the hippocampus bilaterally (Allen, et al. 2006) that was relatively focal. The etiology of 1951 and 2308 was herpes simplex encephalitis, and both had large temporal lobe lesions typical of post-encephalitic patients. Hippocampus was lesioned bilaterally and almost completely for both 1951 and 2308, and significant damage to the medial, ventral, and lateral temporal lobes was evident. In 1951, lesion extent was greater on the left, while in 2308 lesion extent was somewhat greater on the right, in both cases extending to the temporal pole in the more-affected hemisphere. Three of the patients reported here have been described previously in greater detail (Cavaco, et al. 2012, Feinstein, et al. 2010, Warren, et al. 2012). Patient 2563 had a pacemaker that precluded MRI examination, but analysis of CT data combined with his anoxic etiology and neuropsychological profile strongly suggested focal hippocampal atrophy similar to the anoxic patients in our sample.

Brain-damaged comparison (BDC) participants were recruited to provide a sample of similar age and education to the amnesic patients (Table 1). Exclusion criteria included all previously-outlined criteria for the Registry (see above) and the following: damage (lesion or atrophy) to the hippocampus or medial temporal lobe evident on MRI examination; poor vision as noted at neuropsychological exam; and poor performance on neuropsychological tests of memory that might indicate hippocampal

dysfunction (i.e., >2 standard deviations below normative expectations). Etiologies included stroke (N=6), resection (N=5), and subarachnoid hemorrhage (N=2).

A group of healthy normal comparison (NC) participants was recruited from Iowa City, Urbana-Champaign, and the surrounding communities to match the amnesic patients for age and education (Table 1).

Procedure

As a preliminary screening measure, all participants completed the 24-plate Ishihara Color Blindness Test (Ishihara. 1962). Participants who missed any items on the color-blindness test (N=1, a BDC participant) completed an alternative task not reported here.

Test sessions, including consent, practice, and the 3 main test blocks, lasted approximately 1 hour. Short breaks were permitted between main test blocks, and the task was reintroduced after each block for the amnesic participants.

Analysis

Participants who completed fewer than 140 trials for any condition (i.e., 1, 3, or 6 items) were excluded from analysis for that condition. Trials with a response time longer than 15 s were excluded. Four NC participants did not complete at least 140 1-item trials, and so only data from their 3- and 6-item trials were included (degrees of freedom for all inferential tests were adjusted accordingly). Although 140 trials per condition was considered sufficient for accurate parameter estimation (Zhang and Luck. 2008) amnesic patients completed additional trials in each condition when possible in order to ensure robust estimates of each patient's response parameters despite our relatively small sample. Amnesic patients completed the following number of trials: 1846, 900 trials (300 per condition); 1951, 900 trials (300 per condition); 2308, 1745 trials (approximately 580 per condition); 2363, 1078 trials (approximately 360 per condition); 2563, 450 trials (150 per condition). We addressed potential concerns about the effects of additional practice on the performance of amnesic patients by separately estimating parameters for the

first and last 150 trials per condition completed by each patient (2563 was excluded because he completed only 150 trials per condition) and comparing the resulting parameter estimates.

Parameter estimation: We estimated the quantity and quality of the information in memory based on response distributions by using the methodology of Zhang & Luck (2008). We assumed that participant responses were divided into two discrete categories: guessing responses in which the color of the probe was not available at test (due to forgetting or inattention); and knowledge-guided responses in which the color of the probe was remembered. Guessing responses were assumed to be uniformly distributed across the response space. Knowledge-guided responses were assumed to be distributed around the studied color of the probe according to a Von Mises distribution, which is conceptually similar to the circular normal distribution. The Von Mises distribution is centered on a mean (μ), and the concentration of the distribution is indexed by κ. Here, the definition of concentration is the inverse of deviation; therefore more concentrated distributions associated with larger values of κ deviate less from the mean value μ.

We estimated three parameters for each participant: p_r, the probability that the probe item was in memory at test (i.e., the *quantity* of information in memory); μ, the mean of the response distribution; and κ, the concentration of the response distribution around μ (i.e., the *quality* of information in memory). The log likelihood function for the Von Mises distribution is given by Equation 1:

$$L(x;\theta) = \sum_{n=1}^{j} \ln \left(\frac{1}{2\pi \cdot I_0(\kappa)} \cdot e^{\kappa \cdot \cos(\theta_j - \mu)} \right)$$

This describes the log likelihood given observations θ of a Von Mises distribution centered on μ and with concentration of κ; I_0 is the zeroth-order Bessel function. We used maximum likelihood estimation (MLE) (Myung. 2003) to fit a mixture of Von Mises and uniform distributions given by Equation 2:

$$L(x;\theta) = \sum_{n=1}^{j} \ln \left(\frac{p_r}{2\pi \cdot I_0(\kappa)} \cdot e^{\kappa \cdot \cos(\theta_j - \mu)} + \frac{1 - p_r}{2\pi \cdot I_0(0)} \cdot e^{0 \cdot \cos(\theta_j - \mu)} \right)$$

This equation follows from the previous description of the log likelihood function of a single Von Mises distribution given above, but gives the log likelihood for a mixture of two Von Mises distributions: the first represents memory-guided responses, has a concentration equal to κ, and occurs with probability p_r; the second represents guesses, has a concentration of 0 (which describes a uniform distribution on [0,2π]), and occurs with complementary probability $1 - p_r$. This approach was adapted from previous reports (Grimshaw, et al. 2001).

Our MLE approach for each participant's per-condition data was as follows. First, the likelihood function for the mixture distribution described in Equation 2 was submitted to a MLE routine (nlm) included with R software (version 3.0.2). The first successful estimate of 1000 MLE attempts with acceptable parameter values (i.e., 0<=p_r<=1, κ>=0, 0<=μ<=2π) was used and the log-likelihood associated with the fitted parameters was recorded. Second, for the 1-item condition only, the same procedure was repeated using the likelihood function for the Von Mises distribution described in Equation 2; this measure was taken to address problems with fitting the mixture model when p_r asymptotically approached 1. As before, the first successful estimate from 1000 MLE attempts was accepted as correct. Third, the MLE routine was applied to the same data a third time using the likelihood function for a uniform distribution (i.e., the distribution that would be expected based on guessing alone). Using the log-likelihoods from the fitted models, the Aikake information criterion (AIC) (Akaike. 1974) was calculated to evaluate the quality of each fit. Finally, the model that provided the best fit to the observed data (i.e., the model associated with the smallest AIC value) was selected and the fitted values associated with each parameter were recorded. If the simple Von Mises model provided the best fit in the 1-item condition, values for μ and κ were recorded as calculated while p_r was set to 1 to indicate that the remembered item was always available (this was the case for 16 of 33

participants). We note that alternative accounts of visual representations have suggested that simple, non-mixture distributions improve model fits (Bays and Husain. 2008), but we found that the mixture model provided a better fit than the simpler Von Mises model in a substantial majority of our data for the 3- and 6-item conditions (3 items, 36 of 37 participants; 6 items, 33 of 37 participants). For this reason, we used the mixture model for all data from those conditions.

Alternative parameter estimation: An alternative form of the mixture model suggested by Zhang and Luck (Zhang and Luck. 2008) has been suggested by other authors (Bays, et al. 2009). Conceptually, this model differs from the model of Zhang and Luck by acknowledging that some responses may be incorrectly driven by non-target stimuli due to spatial confusion. The potential influence of non-target items is modeled by including an additional model parameter to estimate the probability that the response on a given trial was driven by a non-target item instead of the target. In total, four parameters are proposed: p_T, the probability that a response was driven by a target representation; p_{NT}, the probability that a response was driven by a non-target representation; p_U, the probability that the response was a guess uniformly distributed across the response space; and κ, the concentration of the response distribution (which is assumed to be constant for target and non-target responses). Notably, the three probability parameters must sum to 1, meaning that the model effectively assumes 3 parameters.

We considered this model in addition to the Zhang and Luck model. Mathematically, the model is described in Equation 3:

$$L(x;\theta) = \sum_{n=1}^{j} \ln \left(\frac{p_T}{2\pi \cdot I_0(\kappa)} \cdot e^{\kappa \cdot \cos(\theta_j)} + \frac{1 - p_{NT} - p_T}{2\pi \cdot I_0(0)} \cdot e^{0 \cdot \cos(\theta_j)} + \frac{p_{NT}}{N_{NT}} \sum_{n=1}^{i} \frac{1}{2\pi \cdot I_0(\kappa)} \cdot e^{\kappa \cdot \cos(\theta_{ij})} \right)$$

Equation 3 is very similar to Equation 2 with two significant additions: first, the addition of the p_{NT} parameter measuring the probability of any non-target item driving the response; and second, the corresponding addition of a second summation that reflects the correspondence between N_{NT} non-

target items and the response. Parameter estimates were obtained using a MATLAB program provided by the originating authors (Bays., 2014).

Response time: Our methodology permitted the measurement of response time (RT), but RT was not a dependent variable of primary interest and was analyzed to address a potential confounds to the main analysis. We also analyzed normalized response time (RT_Z). RT was measured as time from test display onset to time of mouse-click response. In the main RT analysis, all responses longer than 15 s or shorter than 0.25 s were discarded prior to analysis (1.110% of all data). In a supplementary analysis intended to control for between-participant differences in response speed, RT data from each participant were Z-transformed (yielding RT_Z) using the following procedure (Faust, et al. 1999). Again, all responses longer than 15 s or shorter than 0.25 s were discarded; next, mean and standard deviation RT were calculated for each participant; and finally, RT_z was calculated for each RT value by subtracting the participant's mean RT value and dividing by the standard deviation of the participant's RT. After calculating RT_z for each RT value, RT_z scores greater than 3 were discarded (a further 1.548% of all data).

Supplemental Results

Additional analysis of main dependent variables

Probability of items being present in memory, supplemental analysis based on neuroanatomy: To evaluate how MTL and hippocampus contributed to the brief representation of information on-line, we conducted a supplemental analysis concentrating on the four patients with MRI confirmed hippocampal atrophy or lesion (i.e., excluding patient 2563). This supplemental analysis showed that the same reduced p_r was reliable in the subgroup of amnesic patients with MRI-confirmed lesion or atrophy of the hippocampus (2563 was excluded because MRI exam was not possible). Numerically, the overall pattern of effects was very similar although the effect of group membership was marginal [$F(2,33)=3.2816$, $p=0.050$]. Planned comparisons showed that in many cases the between-group differences were similar to the previous whole-group analyses, particularly in the 3-item condition.

Specifically, amnesic patients with MRI-confirmed bilateral hippocampal damage had significantly lower values of p_r than the NC group for presentations of 1 item [T(17)=2.142, p=0.047, p_{perm}=0.036, d_{unb}=1.152] and 3 items [T(21)=2.360, p=0.028, p_{perm}=0.017, d_{unb}=1.251], and significantly lower values of p_r than the BDC group for presentations of 3 items [T(15)=2.436, p=0.028, p_{perm}=0.017, d_{unb}=1.322] with marginally lower values of p_r for presentations of 1 item [T(15)=1.857, p=0.083, p_{perm}=0.063, d_{unb}=1.008]. Notably, evidence for a deficit in the amnesic group in the 6-item condition was mixed: permutation tests and large effect sizes suggested that the differences were statistically significant (NC vs. amnesic, p_{perm}=0.041, d_{unb}=0.903; BDC vs. amnesic, p_{perm}=0.049, d_{unb}=0.921), while standard parametric tests did not (each T<1.8, each p>0.10).

Probability of items being present in memory, supplemental commentary on individual differences: As in most neuropsychological investigations of small, relatively rare patient populations, there was some variability in the performance of the amnesic patients (see Table S3). Notably, for the 3- and 6-item conditions in which the amnesic group as a whole was significantly impaired, patient 1846 performed near the normal mean. The reason for patient 1846's relatively normal performance is not obvious. All amnesic patients (and more generally, all participants) had normal color vision as assessed with the Ishihara Color Blindness Test, ruling out color blindness as an explanation. All amnesic patients also had bilateral damage to the hippocampus, and two other anoxic amnesic patients had focal hippocampal damage (MRI-confirmed for 2363), limiting any potential confounds of neuroanatomy. Two possible explanatory factors could be age and sex. Regarding age, 1846 was the youngest of the amnesic patients whom we tested by 7-10 years and age may influence performance on this task. Zhang and Luck (2008) tested a group (N=8) of 18-35 year-old participants whose performance yielded larger estimates of p_r than our somewhat older NC and BDC groups in the 3-item condition (young>older, $\Delta p_r \geq 0.06$). These differences could be attributable to age, although we note that our implementation of the Zhang and Luck (2008) task was independent and therefore may not yield identical performance.

Regarding sex, exploratory analyses not presented in our manuscript showed modest numerical advantages for female comparison participants over male comparison participants for p_r in the 3- and 6-item conditions (F>M, $\Delta p_r \geq 0.06$). It is possible that this difference could be attributable to sex differences in color lexicons (Nowaczyk. 1982), color perception (Bimler, et al. 2004), color memory (Pérez-Carpinell, et al. 1998), or greater prevalence of more than 3 retinal photopigments (Jameson, et al. 2001), but those explanations are beyond the scope of our investigation. In summary, we speculate that influences of age, sex, or an interaction of these factors may have contributed to 1846's relatively normal performance.

Center of response distributions: Response distributions were centered on the correct response for all groups, and did not differ between groups or with the number of items presented. No significant differences were observed between groups [$F(2,34)=1.996$, $p=0.152$], between numbers of items presented [$F(2,64)=0.556$, $p=0.576$], or in the interaction of these factors [$F(4,64)=0.613$, $p=0.655$]. The intercept term was significantly different than 0 [$F(1,64)=4.638$, $p=0.035$] which suggested a consistent but small bias for all groups to respond slightly clockwise (0.025 radians or 1.430˚) from the target position. Because each unique color band on the colored ring occupied 2˚ this 1.430˚ response error fell within the acceptable response range for any given color, and we did not consider this difference further.

Evidence for practice effects in amnesic patients: We found no evidence that additional exposure to test materials significantly affected the performance of amnesic patients. Four of the five amnesic patients (excluding 2563) completed at least 300 trials per condition (versus 150 trials per condition for comparison participants) with the goal of better estimating the parameters of their response distributions. We addressed potential learning effects due to additional exposure by separately estimating parameters for the response distributions in the first and last 150 trials per condition (1-, 3-, and 6-item) collected from each amnesic patient. Parameter estimates of p_r and κ from the first and

last 150 trials for each condition were compared using parametric paired T tests and non-parametric Wilcox tests. Estimated p_r differed no more than marginally between the first and last 150 trials [1-item: T(3)=2.963, p=0.059, Wilcox Z=0.592, p=0.554; 3-item: T(3)=0.657, p=0.558, Wilcox Z=0, p=1; 6-item: T(3)=0.971, p=0.403, Wilcox Z=0.866, p=0.387]. Similarly, estimated κ differed no more than marginally between samples [1-item: T(3)=0.097, p=0.929, Wilcox Z=0.577, p=0.563; 3-item: T(3)=1.091, p=0.355, Wilcox Z=1.443, p=0.149; 6-item: T(3)=0.888, p=0.440, Wilcox Z=0, p=1]. Both p_r and κ were marginally greater in the 1-item condition for the last 150 trials, but the numerical differences were relatively small ($p_{r\,\text{first}}$=0.924 vs. $p_{r\,\text{last}}$=0.935; κ_{first}=8.723 vs. κ_{last}=8.877). No other differences approached statistical significance.

Response time

Response time, raw: Response time (RT) increased with the number of items presented and was greater for the amnesic group than for the NC and BDC groups (Table S2). The number of items presented was generally associated with increased RT for all groups [F(2,64)=38.727, p<0.001]. Within-group planned comparisons showed that the NC and BDC groups responded more quickly when 1 item was presented vs. 3 items [NC, T(32)=2.207, p=0.035; BDC, T(24)=2.923, p=0.007] while amnesic patients did not [T(8)=1.287, p=0.234]. No group responded more quickly when 3 vs. 6 items were presented [each T<1, each p>0.30]. Additionally, the amnesic group responded more slowly overall than the NC and BDC groups [F(2,34)=12.274, p=0.001]. Between-group planned comparisons within levels of number of items presented revealed that this effect was ubiquitous and significant for 1 item [NC, T(18)=2.306, p=0.005; BDC, T(16)=4.219, p=0.001], 3 items [NC, T(22)=4.825, p<0.001; BDC, T(16)=2.909, p=0.010], and 6 items [NC, T(22)=5.468, p<0.001; BDC, T(16)=2.519, p=0.023].

Response time, normalized: Normalized RT (characterized as RT_z) generally increased with the number of stimuli in a display, but there were no between-group differences in RT_z (Table S2). Group membership

did not significantly influence RT_z [F(2,34)=0.158, p=0.854], and there was no interaction of group membership with the number of items presented [F(4,64)=1.083, p=0.373]. The number of items presented did have a significant influence on RT_z [F(2,64)=146.281, p<0.001] that reflected a monotonic increase in RT_z from 1 item to 3 items to 6 items for all groups. For the amnesic and BDC groups, differences between each level of the number of items presented were all significant in planned comparisons [1 vs. 3 items, each T>4.9, each p<0.0025; 3 vs. 6 items, each T>2.7, each p<0.05]. For the NC group, RT_z was least for 1 item and significantly greater for 3 items [T(32)=7.909, p<0.001], while RT_z for 3 and 6 items was similar [T(36)=1.052, p=0.300]. The observed patterns of RT_z showed that when individual differences in RT were normalized, the effects of presenting more items were generally similar for all groups.

Relationship between probability of memory and raw response time

The amnesic group had a significantly reduced probability of remembering items at test (p_r) relative to the NC and BDC groups, but also had longer RT. Although our methodology could not directly address a causal relationship between p_r and RT, we analyzed the relationship between p_r and RT in the NC and BDC groups when 3 items were presented. First, we evaluated within-group correlations between p_r and RT for the NC and BDC groups. Second, we used a split-half analysis to directly compare fast and slow subgroups of the NC and BDC groups to determine whether faster and slower responding were significantly related to better performance in those groups. Neither analysis found evidence that slower responding in comparison groups was significantly related to smaller values of p_r on the timescales of our task (see also Zhang and Luck. 2009). These results cast doubt on the proposition that longer RT produced a selective disadvantage for amnesic patients.

Correlation of p_r and raw response time: If longer response times were associated with smaller values of p_r, then a significant negative correlation between p_r and RT would be expected for comparison

participants. We did not observe a significant correlation between p_r and RT in the NC group, the BDC group, or combination of the two. Specifically, RT was not significantly correlated with p_r for the NC group (r=-0.336, T(17)=1.469, p=0.160), for the BDC group (r=-0.208, T(11)=0.7043, p=0.496), or for the combined comparison group (r=-0.299, T(30)=1.717, p=0.096).

Comparing p_r in fast and slow comparison participants: If longer response times were associated with smaller values of p_r, then participants who responded faster on average would be expected to have higher values of p_r than participants who responded slower on average. We did not observe significant differences in p_r in fast vs. slow NC participants (N=9 each), BDC participants (N=6 each), or a combination of the two groups (N=16 each). Specifically, fast vs. slow NC participants did not differ significantly [fast NC p_r=0.801(0.039); slow NC p_r=0.763(0.054); T(16)=0.565, p=0.580], fast vs. slow BDC participants did not differ significantly [fast BDC p_r=0.749(0.032); slow BDC p_r=0.717(0.033); T(10)=0.676, p=0.515], and fast vs. slow comparison participants drawn from both groups did not differ significantly [fast comparisons p_r=0.778(0.028); slow comparisons p_r=0.729(0.032); T(30)=1.156, p=0.257].

Alternative model parameters

We conducted an exploratory analysis of our behavioral data using an alternative model of response distributions. The alternative model described by Bays et al. (2009) makes different assumptions than the Zhang and Luck (2008) model about how target and non-target items contribute to the response distribution. The model describes the response distribution using 4 parameters (but note that any two of the three probability parameters constrain the third): the probability of the target item driving the response (p_T); the probability of a non-target item driving the response (p_{NT}); the probability of guessing (p_U); and the concentration of the response distribution (κ). Key predictions by the originators of this model (Bays, et al. 2009) are that with increasing numbers of items: p_T should decrease; p_{NT}

should increase; p_U should increase; and κ should decrease (i.e., the response distribution should become less concentrated and spread wider). Overall model fit and the parameters are considered below (see also Table S4), and agreement with specific predictions is discussed afterward.

Model fit: Before analyzing individual parameters, we compared the quality of model fit between the Bays et al. model and the Zhang and Luck model. Model fit was generally better for the Zhang and Luck model. Comparisons were conducted using AIC values of the fitted models for each participant in each multi-item condition. The Zhang and Luck model had a smaller AIC value (corresponding to better fit) for 24 of 37 participants in the 3-item condition and 20 of 37 participants in the 6-item condition. This pattern of results implied that the Zhang and Luck model predicted the observed data as well or better than the Bays et al. model for most participants. While empirical measures of model fit such as AIC are not always sufficient to evaluate whether the theoretical and conceptual predictions of a model justify its selection, it is notable that the consideration of non-target item responses by the Bays et al. model did not substantially improve fit. Additionally, the Zhang and Luck model was (effectively) penalized in the AIC comparison for incorporating a distribution mean parameter (μ) which the Bays et al. model assumes to be 0 (nb. our empirical results agree with this assumption: see Supplemental Results, Center of response distributions).

Probability of target item driving response (p_T): The amnesic group showed reduced values of p_T overall relative to the comparison groups [$F(2,35)=4.865$, $p=0.014$]. Specifically, the amnesic group had a significantly lower probability of basing a response on the target item than either comparison group in the 3-item condition, and also had significantly lower p_T values than the NC group in the 6-item condition (Table S4). Evidence for reduced amnesic p_T versus the BDC group in the 6-item condition was mixed depending on the specific test used ($p=0.118$, $p_{perm}=0.024$; see Table S4), but the effect size was moderately large ($d_{unb}=0.820$). Also, all groups had significantly lower values of p_T in the 6-item

condition than the 3-item condition [amnesic group, T(4)=6.443, p=0.002; NC group, T(18)=10.489, p<0.001; BDC group, T(13)=5.443, p<0.001].

Probability of non-target item driving response (p_{NT}): There were no significant between-group differences in the probability of a response being based on a non-target item [F(2,35)=0.182, p=0.834], and there was no uniform pattern to the group estimates of p_{NT} between conditions (Table S4). For example, the amnesic group had a smaller p_{NT} value in the 6-item condition than the 3-item condition while the comparison groups showed the opposite pattern. Effect sizes for between-group contrasts were uniformly smaller for p_{NT} than for p_T and p_U, and the directionality of group differences changed between conditions as well. Within groups, there were no significant differences between the 3-item and 6-item conditions [amnesic group, T(4)=1.695, p=0.162; NC group, T(18)=1.341, p<0.197; BDC group, T(13)=1.738, p=0.106]. However, all three groups had values of p_{NT} that were significantly greater than zero in both the 3-item and 6-item conditions (one sample T tests vs. 0, each p<0.05).

Probability of guess response (p_U): Overall, the amnesic group showed increased guessing responses versus the comparison groups [F(2,35)=3.369, p=0.046]. In support of this, the amnesic group had numerically larger values of p_U than either comparison group in both the 3-item and 6-item conditions and effect sizes for between-group contrasts were relatively large (each d_{unb}>0.82) (Table S4). Findings from specific planned comparisons were more nuanced. Standard T tests suggested that the amnesic group was had marginally greater values of p_U than both comparison groups, while permutation tests suggested that these differences were significant. Within-group contrasts showed that all groups had larger values of p_U in the 6-item condition [amnesic group, T(4)=6.819, p=0.002; NC group, T(18)=8.148, p<0.001; BDC group, T(13)=5.296, p=0.001], likely reflecting increasing guessing in the presence of additional items.

Concentration of response distribution (κ): There were no significant between-group differences in the concentration of response distributions [F(2,35)=0.123, p=0.884]. Effect sizes for between-group

contrasts were very small (i.e., <0.25) relative to effect sizes for p_T, p_{NT}, and p_U. Within groups, there were no significant differences in κ between the 3-item and 6-item conditions [amnesic group, T(4)=0.814, p=0.461; NC group, T(18)=1.083, p=0.293; BDC group, T(13)=0.331, p=0.746].

Summary: The alternative model of Bays et al. (2009) did not substantially improve on the Zhang and Luck model when fit to our observations, but the estimated parameter values generally agreed with our main findings. According to those values, the amnesic group showed a reduced probability of basing their response on the target item relative to both comparison groups, and the amnesic group also showed an increased probability of guessing. This evidence for increased guessing responses by the amnesic group aligned closely with our main findings (see Figure 2A). Meanwhile, there were no significant between-group differences in responses to non-target items, although all groups showed a significant probability of responding to a non-target item versus floor. We also did not observe reduced concentration of response distributions (κ) between groups or with larger numbers of items, also much like our main findings (see Figure 2B, 3 items vs. 6 items). While these findings matched the predictions of the Zhang and Luck model quite well, support for the predictions of the Bays et al. model was very limited.

We note that our analysis using the Bays et al. model was exploratory and our design was not optimized for discriminating between the two models. It remains possible that amnesic patients are more likely to incorrectly respond to a non-target item in this paradigm, which we would attribute to an impaired ability to rapidly form relations between a spatial location and a color. However, our current findings indicate that guessing responses play a larger role in amnesic performance than spatial errors. Future work could be tailored to address this question with greater specificity.

Supplemental Discussion

Performance declined for all groups when more items were presented, but we do not believe that floor effects drove our findings for any parameter in the 3- or 6-item conditions. First, we collected

additional data from amnesic participants in order to address concerns of insufficient data for parameter fitting. Second, memory in the 3-item condition was relatively good for all groups, as >56% of stimuli were remembered by all groups on average. Considering the large number of trials collected from amnesic patients, this means that parameter estimates for the memory-guided response distributions that included the accuracy parameter (κ) arose from more than 150 trials per patient. Parameter estimates in the 6-item condition were based on fewer memory-guided trials and were more variable as a result (see Figure 2B and S1B). Third, the overall pattern of our findings across the 1-, 3-, and 6-item conditions was similar to previous reports (e.g., Zhang & Luck, 2008), with higher values of κ in the 1-item condition and lower values of κ in the 3- and 6-item conditions that did not differ from one another. Fourth, the lower limit of κ in the Zhang and Luck model model is 0 (which reflects a uniform distribution with no peak), while our lowest group mean values were approximately 5, providing substantial separation between observed values and the κ parameter's minimum value. Thus, we do not believe that floor effects played a significant role in our findings in any condition, and are very confident that this was not the case in the 3-item condition.

Supplemental Figure Captions

Figure S1: Fitted models for each group and aggregate response accuracy data. Purple lines represent mixture distributions as in Figure 1A. A) Estimated response distributions for the NC, BDC, and amnesic groups in the 1-, 3-, and 6-item conditions are presented as purple lines (see also Figure 1B). The reduced accuracy of the amnesic group is reflected in a lower peak and higher tails. B) Fitted models for each group in the 1-, 3-, and 6-item conditions (as in A) plotted with point clouds reflecting aggregated response frequency for all 180 possible levels of accuracy relative to the target value on each trial. In each case, the estimated response distribution appears to accurately capture key characteristics of the

underlying response distribution, as reflected in the proportion of aggregate variance accounted for by the estimated response distribution (top right corner of each plot).

Supplemental References

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control***19:** 716-723.

Allen, J.S., Tranel, D., Bruss, J. and Damasio, H. 2006. Correlations between regional brain volumes and memory performance in anoxia. *J Clin Exp Neuropsychol***28:** 457-476.

Bays, P.M. Action & cognition lab | paul bays. **2014:** http://paulbays.com/code/JV10/index.php.

Bays, P.M., Catalao, R.F. and Husain, M. 2009. The precision of visual working memory is set by allocation of a shared resource. *J Vis***9:** 7.1-11.

Bays, P.M. and Husain, M. 2008. Dynamic shifts of limited working memory resources in human vision. *Science***321:** 851-854.

Bimler, D.L., Kirkland, J. and Jameson, K.A. 2004. Quantifying variations in personal color spaces: Are there sex differences in color vision? *Color Research & Application***29:** 128-134.

Cavaco, S., Feinstein, J.S., van Twillert, H. and Tranel, D. 2012. Musical memory in a patient with severe anterograde amnesia. *J Clin Exp Neuropsychol***34:** 1089-1100.

Faust, M.E., Balota, D.A., Spieler, D.H. and Ferraro, F.R. 1999. Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychol Bull***125:** 777-799.

Feinstein, J.S., Rudrauf, D., Khalsa, S.S., Cassell, M.D., Bruss, J., Grabowski, T.J. and Tranel, D. 2010. Bilateral limbic system destruction in man. *J Clin Exp Neuropsychol***32:** 88-106.

Grimshaw, S.D., Whiting, D.G. and Morris, T.H. 2001. Likelihood ratio tests for a mixture of two von mises distributions. *Biometrics***57:** 260-265.

Ed. Ishihara, S. 1962. Tests for colour-blindness: 24 plates. Kanehara Shuppan Co., Ltd., Tokyo, Japan.

Jameson, K.A., Highnote, S.M. and Wasserman, L.M. 2001. Richer color experience in observers with multiple photopigment opsin genes. *Psychon Bull Rev***8:** 244-261.

Myung, I. 2003. Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology***47:** 90-100.

Nowaczyk, R.H. 1982. Sex-related differences in the color lexicon. *Language and Speech***25:** 257-265.

Pérez-Carpinell, J., Baldoví, R., de Fez, M.D. and Castro, J. 1998. Color memory matching: Time effect and other factors. *Color Research & Application***23:** 234-247.

Warren, D.E., Duff, M.C., Magnotta, V., Capizzano, A.A., Cassell, M.D. and Tranel, D. 2012. Long-term neuropsychological, neuroanatomical, and life outcome in hippocampal amnesia. *Clin Neuropsychol***26:** 335-369.

Zhang, W. and Luck, S.J. 2009. Sudden death and gradual decay in visual working memory. *Psychol Sci***20:** 423-428.

Zhang, W. and Luck, S.J. 2008. Discrete fixed-resolution representations in visual working memory. *Nature***453:** 233-235.