Atten Percept Psychophys. Author manuscript; available in PMC 2014 September 03.

Published in final edited form as:
Atten Percept Psychophys. 2012 January ; 74(1): 146–162. doi:10.3758/s13414-011-0218-3.

Orienting attention to locations in mental representations

Duncan Edward Astle[1,2], **Jennifer Summerfield**[2], **Ivan Griffin**[2], and **Anna Christina Nobre**[2]

[1]Department of Psychology, Royal Holloway, University of London, UK.

[2]Department of Experimental Psychology, University of Oxford, UK.

Abstract

Many cognitive processes depend on our ability to hold information in mind, often well beyond the offset of the original sensory input. The capacity of this ‘visual short-term memory’ (VSTM) is limited to around three to four items. Recent research has demonstrated that the content of VSTM can be modulated by top-down attentional biases. This has been demonstrated using retrodictive spatial cues, termed ‘retro-cues’, which orient participants’ attention to spatial locations within VSTM. In the current paper, we tested whether the use of these cues is modulated by memory load and cue delay. There are a number of important conclusions: i) top-down biases can operate upon very brief iconic traces as well as older VSTM representations (Experiment 1); ii) when operating within capacity, subjects use the cue to prioritize where they initiate their memory search, rather than to discard un-cued items (Experiments 2 and 3); iii) when capacity is exceeded there is little benefit to top-down biasing relative to a neutral condition, however, unattended items are lost, with there being a substantial cost of invalid spatial cueing (Experiment 3); iv) these costs and benefits of orienting spatial attention differ across iconic memory and VSTM representations when VSTM capacity is exceeded (Experiment 4).

Keywords

VSTM; visual short-term memory; visual working memory; iconic memory; spatial attention; partial report

1. Introduction

Attention enables us to bias the processing of incoming perceptual information, favouring those aspects of the input most relevant to the task at hand [Posner, 1980]. In addition to operating upon perceptual representations, recent studies have demonstrated that attention can bias the contents of visual short-term memory [VSTM; Astle et al, 2009; Griffin & Nobre, 2003; Landman, Spekreijse & Lamme, 2003; Lepsien, Griffin, Devlin, & Nobre, 2005; Lepsien & Nobre, 2006; Matsukura, Luck and Vecera, 2007; Makovski, Sussman and Jiang, 2008; Sligte et al. 2008, 2009, 2010].

Address for correspondence: Anna Christina Nobre University of Oxford Department of Experimental Psychology South Parks Road Oxford, OX1 3UD, UK Facsimile: 44-1865-310447 Telephone: 44-1865-271324 kia.nobre@psy.ox.ac.uk.

Whilst much is known about how we bias aspects of our perceptual representation, how the contents of memory can be biased is only beginning to be explored. Recent studies have used attention-directing retro-cues to demonstrate that the contents of VSTM are available for top-down biasing. Retro-cues are presented after the to-be-remembered arrays of items, and predict which location or item from within the remembered array will be relevant to a subsequent decision based on the memory representation, such as judging whether a probe item was contained within the array. Retro-cues give participants time to orient their attention to the cued location prior to the probe stimulus appearing. Retro-cues, therefore differ from post-cues, which directly act as a probe for accessing a specific aspect of the remembered array. Post-cues give no time to orient; subjects are instructed to respond immediately. Sligte and colleagues [2008] compared performance on a post-cue with that on a retro-cue. They demonstrated that capacity estimates more than doubled when participants received a retro-cue compared to a post-cue, even though the interval between the array and the cue was the same in both cases, and the actual delay until the probe item was longer in the retro-cue condition. The authors argued that there existed a store of many 'fragile' VSTM representations, which were wiped by the onset of the post-cue, but which were boosted by the attention resulting from the retro-cue, meaning that they could survive the onset of the probe.

The current set of experiments seeks to extend our understanding of how these retro-cues operate upon memory representations. Work so far has ruled out a series of uninteresting artifactual explanations for some of these effects, for instance: speed-accuracy trade-offs [Griffin & Nobre, 2003; Lepsien et al. 2005], response biases [Griffin & Nobre, 2003], eye-movements [Griffin & Nobre, 2003; Matsukura, Luck & Vecera, 2007], or articulation [Makovski, Shin & Jiang, 2006; Makovski, Sussman & Jiang, 2008; Matsukura et al. 2007]. Other, more interesting possibilities remain: the retro-cue might enable participants to enhance the active maintenance of cued items and/or suppress the active maintenance of uncued items; retro-cues might enable participants to insulate particular items from decay or interference; or retro-cues might enable subjects to prioritise where they start their memory search. (It is important to note that these possibilities need not be mutually exclusive).

In the current set of experiments we used valid, invalid and neutral retro-cues. As with previous experiments, the valid retro-cues enabled us to explore the advantage of committing attention to a stored item, relative to not committing attention to any one item (i.e. relative to a neutral cue baseline, e.g. Griffin and Nobre, 2003). Furthermore, using invalid cues enabled us to test the fate of those items that were un-cued, by occasionally probing those items on a subset of trials. One possibility is that subjects use the cue to select a single item and discard un-cued items from memory (maybe by allowing them to degrade, see Matsukura et al. 2007), meaning that invalid cues would have a catastrophic effect on accuracy measures. The un-cued items would be lost and be unavailable for re-inspection at probe onset. By contrast, if subjects merely use the retro-cue to prioritise where to start their VSTM search at probe onset, rather than actually discard some items, then invalid cues will have relatively little effect on accuracy measures. Those un-cued items would still be maintained in VSTM, just as well as the cued items, but subjects would initiate their VSTM search in the wrong location; whilst their responses would be slowed, little accuracy cost, relative to neutral cues, should be observed. In short, manipulating cue validity in this way enables us to ask important questions as to how these attentional mechanisms operate upon the contents of memory to bring about validity benefits and/or invalidity costs. This distinction is similar to the distinction between ‘protection’ and ‘prioritisation’ given by Matsukura et al. (2007); the difference being that we are suggesting that a pure prioritisation account, in which cued and uncued items are equally available at probe onset, would predict cue validity effects on RT but not on accuracy.

There are various ways of exploring the effect of retro-cueing on the uncued items. Matsukura and colleagues (2007) used a double cue paradigm. Participants were retro-cued to a subset of memory array shapes, and then, on a minority of trials, retro-cued to the other half of the shapes – i.e. those that were previously uncued. This initial cue acted like an invalid cue, cueing subjects to a subset of items that would not needed at probe onset. Accordingly, accuracy rates were much lower when subjects following these alternating double cue trials. Having discounted various alternative accounts, the authors concluded that subjects do not merely use the retro-cues to prioritise where they start searching their memory; rather, the effect of the retro-cues is that un-cued items are lost from VSTM. This could be either because subjects simply stop maintaining / insulating them from decay, or actively suppress them in order to remove them VSTM. Either way, it is clear that the retro-cue does more than prioritise where to initiate memory search. An alternative way of testing the availability of uncued items is to use the single-cue paradigm, and, on a minority of trials probe the uncued item/s. This approach has the advantage of removing the additional passage of time from conditions for assessing the state of uncued items. Matsukura et al. (2007) demonstrated that the cost of an invalid single cue is very similar to the cost of an initial invalid cue in the double cue paradigm. If subjects use retro-cues to protect certain items from being lost (or to lose uncued items intentionally) then subjects should perform poorly when one subsequently cues (double cue) or probes (invalid single cue) those previously uncued items. However, the relative costs and benefits of retro-cueing, and thus how subjects use the retro-cue, may change as a function of various factors, such as cue-delay and memory load. That is, subjects may use the cues differently depending upon these factors. The modulation of the benefits and costs of retro-cueing, by these two factors, is the principle question of the four experiments herein.

1.1.The effect of cue latency on retro-cue costs and benefits

Sperling (1960) used a partial report paradigm, to demonstrate that more information is ‘available’ from large arrays than can be reported under usual task conditions. When subjects performed a whole report on an array of visual letters – being asked to report as many items as possible – they were reliably able to retrieve up to around four items. However, when they were selective probed about the contents of only one of the rows according to an auditory cue following the array, the estimate of the number of items maintained over a very brief period was much higher. When the cue was delivered at 150 or 300 ms post array offset, the estimated number of items available was, in some cases, double that of the whole-report condition (the effect was subsequently replicated with visual cues, Averbach and Corriell, 1961). However, when this cue was delayed by 1000 ms, this partial report benefit was lost, with performance being equivalent to the whole-report condition. Sperling concluded that a substantial number of items are held in memory immediately after the offset set of the array; however, within the first second of this maintenance delay all but around four of these are lost [Averback & Coriell, 1961; Dick, 1974; Sperling, 1960]. Of course, the specific array-cue intervals inevitably provide underestimates of the longevity of different types of memory traces, since it will undoubtedly take some time to process and interpret the auditory cues, meaning that the availability of these extra items might persist for slightly longer than Sperling had estimated.

In light of this functional distinction proposed between very brief, iconic memory (Neisser, 1967) and the more durable VSTM, we asked whether top-down attentional biases may work differently upon these types of traces to enhance memory-based performance. We asked whether the relative benefits and costs of retro-cues may differ depending upon whether they operate on this iconic memory trace (IM) or a more durable VSTM representation, this is explored in Experiments 1, 2 and 4.

1.2.The effect of memory load on retro-cue costs and benefits

By contrast with IM, the capacity of VSTM is strictly limited [Bays, Catalao & Husain, 2009; Cowan, 2001; Vogel & Machizawa, 2004; Vogel, Woodman & Luck, 2001; Zhang & Luck, 2008]. For example, people's failure to notice large changes in visual scenes unless attention is specifically directed to them, is usually used to infer some limit to VSTM capacity [O'Regan, Rensink & Clark, 1999; Rensink, O'Regan & Clark, 1997]. In a series of experiments, Luck and colleagues, as well as others, suggested that VSTM can only store approximately four integrated visual items, irrespective of the number of constituent features [Vogel et al. 2001; Lee & Chun, 2001; Luck & Vogel, 1997].

It is possible, therefore, that the relative costs and benefits of spatial retro-cues differ across load. This is explored in Experiments 3 and 4. For instance, when the array size is comfortably within capacity (=< 4 items), subjects may use the spatial retro-cue to prioritise where they initiate their memory search, but maintain all of the items nonetheless, producing little accuracy cost of invalid cueing. Alternatively, when the number of to-be-remembered items exceeds capacity, subjects may use the cue to prioritise which item/s they maintain, discarding those un-cued items, thereby reducing the memory load. In short: the way in which a retro-cue is used may differ depending upon whether or not VSTM capacity is exceeded.

Given what has already been demonstrated [e.g. Sperling, 1960; Sligte et al. 2008], we might expect these two factors of cue-latency and memory load to interact. When cue delays are short, a larger capacity, IM system, is tapped; when cue delays are long, a smaller capacity, VSTM, system is tapped. In our final experiment we tested this, by manipulating both cue SOA and memory load in the same design, to explore the potentially interactive effects of these factors on subjects' use of the retro-cue.

Across the experiments, we report three different dependent measures: RT is an index of the speed with which items can be retrieved correctly, but it is insensitive to the number of items available for retrieval. Given that many of our questions pertain to the availability of items for retrieval, rather than the speed with which they can be retrieved, measures of accuracy might be more important. We present two measures of accuracy: i) d' indexes accuracy, taking into account any response biases; ii) K is also based on participants retrieval accuracy, but takes into account the number of items in the original array [Cowan, 2001].

2.1: Experiment 1: Benefits of spatial orienting to items held in iconic or visual short-term memory

Most previous experiments have investigated the advantages gained from retrospective attentional cueing when the contents of VSTM, upon which those retro-cues operate, have been stored for between 1 and 2.5 seconds [Astle, Nobre & Scerif, 2010; Griffin & Nobre, 2003; Nobre, Griffin & Rao, 2008]. In one recent study [Sligte et al. 2008] retro-cue SOA varied up to 5 seconds, and even at these long retention intervals the retro-cue provided a benefit relative to performance on post-cue trials. However, this previous study did not compare retro-cue performance to that on a neutral baseline, so it could not be determined whether the *benefit* of retro-cues changed as the contents of VSTM decay, relative to a situation in which attention is never committed to any single item. We tested the temporal limits of retrospective attentional cueing, by varying the retention period between the onset of the to-be-remembered array of items and the retrospective cue, and we used the neutral cues as a baseline condition. We varied the SOA between the array and retro-cue using seven intervals between 150 and 9600ms. This also enabled us to address an additional question of whether retro-cues provide a similar degree of benefit when they operate on the contents of iconic memory (IM, e.g. 150ms) or VSTM (e.g. 2040ms). We were keen to test whether there was continuity in the benefit derived from retrospective attentional cueing across these delays. One might predict a decreasing relative benefit of retro-cues with increasing SOA – as the contents of VSTM decay so the stored items are not as available for attentional selection. Alternatively, one might predict that at longer durations VSTM items become too fragile for retrieval via standard serial search mechanisms [Sternberg, 1966], but still available for attentional selection. Moreover, this attentional selection (following a retro-cue) might provide the stability needed for successful retrieval [Sligte et al. 2008]. Were that the case, we might find that the *relative benefit* of valid retro-cues would increase with increasing decay.

In addition to this effect of decay, one might predict that the benefit gained from a valid retro-cue would be different when operating upon the contents of IM, by comparison with VSTM, because of the differing amount and nature of information available for attentional selection in these two forms of storage [Averback & Coriell, 1961; Sperling, 1960].

2.2. Materials and Methods

2.2.1. Participants

The experimental methods in this and all subsequent experiments had ethical approval from the Central University Research Ethics Committee at the University of Oxford, UK. Participants were healthy paid volunteers from the University of Oxford community of students and researchers. All had normal or corrected to normal visual acuity and were right-handed. Twelve participants (age range 20-31 years, 7 females) took part in this experiment.

2.2.2. Stimuli and Task

The task is illustrated in Figure 1. Participants viewed arrays of four differently coloured crosses followed by an informative or neutral cue, and made a delayed decision about the colour of the items in the array. At the end of each trial sequence a probe stimulus appeared, their task was to decide whether the probe stimulus had been in the preceding memory array. The time interval between the onset of the array and cue (stimulus onset asynchrony, SOA) varied between 150 and 9600ms, with the interval doubling successively across seven intervals (i.e. 150, 300, 600ms etc).

Each trial contained the same sequence of events. A square (side length 0.8°, central cue stimulus) appeared at the centre of the screen for a random interval between 600 and 900ms. An array of four differently coloured crosses was then presented for 100ms. The crosses were any four of the following colours: red, blue, green, yellow, orange, cyan, pink, or grey. Each cross was 0.8° visual angle in size, and centred at 3° horizontal and 3° vertical eccentricity. There was then another interval between 50 and 9500ms (150 and 9600ms SOA), after which either an informative or neutral cue was presented. The cue appeared at one of seven possible SOAs with equal probability and in a randomised fashion: 150, 300, 600, 1200, 2400, 4800, or 9600ms. Following the cue, after a random interval between 500 and 1000ms, a coloured cross (any one of the eight possible colours; probe stimulus) appeared at the centre of the screen for 100ms. Participants responded by pressing the left button of a response box with their right-hand index finger if the probe stimulus did appear in the array (“yes” response), and the right button with their right-hand middle finger if the probe stimulus did not appear in the array (“no” response).

An informative cue consisted of two adjacent sides of the central square brightening (forming an arrow). Informative cues occurred in 67% of trials and accurately predicted (with 100% validity) the location where the probe stimulus occurred if it had been present in the array. Neutral cues occurred in 33% of trials. They consisted of the whole square brightening and gave no spatial information about the likely location of the probe.

The probability of correct “yes” or “no” responses was equal; i.e. 50% of the time the probe stimulus was present in the array, and 50% of the time it was not. This was true for both informative and neutral trials, at all SOAs. Both informative and neutral trials, at all SOAs, occurred interspersed in a random order throughout the experiment.

There were 588 trials (392 informative, 196 neutral; 84 at each SOA). Of the informative trials, 196 were target-present valid trials (probe appeared in array at the cued location) and 196 were target-absent trials. Of the neutral trials, 98 contained a probe that was in the array and 98 contained a probe that was not in the array. Within informative and neutral trials, the seven SOAs (150, 300, 600, 1200, 2400, 4800, 9600ms) and the four cue directions (top-left, top-right, bottom-left, bottom-right) occurred equally frequently. There were fourteen blocks of trials in the experiment, plus one additional practice block at the beginning. In all of our analyses, where the assumption of sphericity is violated, the Greenhouse-geisser correction for non-sphericity is applied.

2.2.3. Procedures

Participants were comfortably seated in a dimly illuminated room, facing a computer monitor placed 100cm in front of them. They were informed about the relationship between the cue, array and probe stimuli. They were asked to maintain fixation on a small cross that was continuously present at the centre of the monitor. They were instructed to respond as quickly as possible following probe stimulus onset, whilst avoiding mistakes. No feedback was given during the experiment.

2.3. Results

2.3.1. Accuracy

The accuracy scores were converted into d' scores (the normalized proportion of correct hits minus the normalized proportion of false alarms). This provides an advantage over analyzing the raw accuracy scores in that it controls for any general response bias that might vary as a function of cue-type (neutral versus valid). These were submitted to a two-way ANOVA, with the within-subjects factors of cue-type and SOA. There was a significant effect of SOA on d' scores [F(6,66)=3.008, p=0.012], with there being a significant trend for decreasing d' scores with increasing SOA [linear contrast for SOA, F(1,11)=14.487, p=0.003]. There was a significant effect of cue [F(1,11)=10.906, p=0.007], with d' scores being higher for valid [2.44 ± 0.23 (Std.E.)] than for invalid [2.05 ± 0.31 (Std.E.)] cues. There was no significant interaction between cue and SOA [F(6,66)=0.283, p=0.943].

2.3.2. Reaction Times

Only correct trials were used in the RT analyses. The RT data were submitted to a three-way repeated-measures ANOVA, with the within-subjects factors of SOA, cue-type and presence (probe present or absent). This revealed a significant main effect of cue [F(1,11)=20.174, p=0.001], reflecting quicker responses on informative [525ms ± 24ms (Std.E.)] than neutral trials [630ms ± 20ms (Std.E.)]. A main effect of target presence [F(1,11)=43.172, p<0.001] indicated quicker responses to probes that had been present [578m ± 27ms (Std.E.)] in the array compared with those that had not [676ms ± 22ms (Std.E.)]. There was also a significant main effect of SOA [F(6,66)=11.674, p<0.001]. In a follow-up analysis we established that this effect of SOA did not stem from a significant linear trend of increasing RTs with increasing SOA (as we might expect given the d' scores) [F(1,11)=1.164, p=0.304], but rather from a significant quadratic trend [F(1,11)=35.867, p<0.001]. The pattern of RTs formed a U-shape. Post-hoc contrasts revealed that RTs in trials differed significantly between 600ms and 1200ms [F(1,11)=26.98, p<0.001], between 2400ms and 4800ms [F(1,11)=10.92, p=0.007], and between 4800ms and 9600ms [F(1,11)=18.45, p=0.001]. In short, RTs significantly decreased between 600ms and 1200ms, and then increased again after 2400ms, resulting in a significant quadratic trend. There was also a significant interaction between probe presence and SOA [F(6,66)=2.315, p=0.043]. This was because the effect of SOA was greatest on probe-present trials [F(6,66)=10.571, p<0.001], relative to the effect of SOA on probe-absent trials [F(6,66)=4.707, p<0.001]. The RTs from probe-present trials can be seen in Figure 2.

2.4. Discussion

Experiment 1 explored the relative benefits of retrodictive cues when they operate across SOAs within the IM and VSTM range, even up to ~10 seconds after the offset of the array. To our knowledge this is the first time that this has been done. Behavioural benefits of orienting attention to internal representations were evident over a range of time periods, including SOAs at which representations are thought to be held in IM, as well as much longer SOAs where representations are held in VSTM. The benefit derived from retro-cues relative to neutral cues was largely consistent across the various SOAs. Despite decreasing d’ scores with increasing SOA, the relative d’ benefit of valid retro-cues did not mirror this pattern (or show the opposite pattern).

Recent studies have shown that retro-cueing is effective in the boosting of decayed ‘fragile’ VSTM representations [Sligte et al. 2008], but that the benefit of a retro-cue, relative to performance on a post-cue trial, decreases with increasing SOA up to 4 seconds. After 4 seconds, performance in both the retro- and post-cue conditions appears to decline. Thus our data replicate the finding of decreasing performance with valid retro-cues with increasing decay, but because we compare this decreasing performance with a neutral condition, we can see that the rate of performance decline is roughly equivalent across these two trial types, and not be because retro-cues become gradually less effective per se.

It might appear that our results contradict previous demonstrations that cues are most effective when operating on an IM representation [e.g. Sperling, 1960]. However, these enhanced benefits have been apparent when subjects are shown supra-VSTM-capacity size arrays and are presented with post-cues, which act as an imperative prompt for retrieval. These classic studies have been used to infer the existence of some iconic store with a capacity that exceeds that of VSTM. In our case the number of items does not exceed VSTM capacity and our retro-cues, unlike post-cues, enable preparation for a subsequent imperative stimulus. Thus our demonstration that the cues operate equally well across the different SOAs does not necessarily contradict previous findings. Of course, if more than four items are available at the iconic time range, then we might expect the factors of cue SOA and memory load to interact, with the cue being most effect when operating on supra-VSTM-capacity IM (see Experiment 4).

To summarise briefly: the results of Experiment 1 extend the previous observations of retro-cue benefits [e.g. Griffin and Nobre, 2003; Lepsien et al., 2005; Sligte et al. 2008), by demonstrating that the benefit of committing attention to a particular item, relative to not committing it to any one item, is evident over a very broad time range; searching through representations held as brief IM traces or more durable VSTM representations can be biased with spatial attention, even if they are stored for ~10 seconds. However, the effect that this cueing has on the remaining un-cued items remains unclear, something addressed by Experiment 2.

3.1: Experiment 2: Benefits and costs of orienting spatial attention to items in iconic and visual short-term memory

It is well documented that valid retro-cues confer a substantial retrieval benefit [e.g. Griffin and Nobre, 2003; Landman et al. 2003; Makovski et al. 2008; Nobre et al. 2008], and Experiment 1 explored the temporal constraints of this benefit. However, why is retro-cueing so advantageous? Using invalid cues is one potential way of exploring this [see also Matsukura et al. 2007]. One possibility is that subjects use the cue to *prioritise* their memory search. If this were the case the occasional invalid cue would slow subjects down, as their search would be initiated in the wrong location; this would, however, have relatively little effect on accuracy scores, as the un-cued items would still be available for re-inspection. Alternatively, subjects might use the retro-cue to *discard* un-cued items, essentially reducing the memory load to a single item. Were this the case, subjects' accuracy scores would be detrimentally affected on occasional invalid trials; the un-cued items would no longer be available and subjects would therefore infer that this was a target-absent trial. In short, using invalid trials offers us the opportunity to explore what happens to the un-cued items when subjects oriented their attention within memory.

In Experiment 2, as with the cueing benefit in Experiment 1, we explored how the cost of invalid cueing changed across SOA. Whilst valid retro-cueing benefits might be relatively stable across SOA, invalid retro-cueing costs may not. For instance, if subjects have been invested in maintaining all four array items for a long period of time (e.g. ~ 10 seconds), then it may be easier simply to keep maintaining the items but prioritise the location that they start their memory search. By contrast, if they have only been maintaining the items for a very brief period of time (e.g. 150 ms), then subjects may opt to use the cue to discard the un-cued items, and save themselves the effort of maintaining them all. In short: how subjects use the retro-cue, and thus what happens to the un-cued array items, may change depending upon how late the retro-cue is presented. This is explored in Experiment 2.

3.2. Materials and Methods

Unless stated otherwise, the methods used in Experiment 2 are identical to those in Experiment 1.

3.2.1. Participants

Twelve participants (age range 20-32, 7 females) took part in the experiment. All were right-handed and had normal or corrected-to-normal vision.

3.2.2. Stimuli and Task

In contrast to Experiment 1, informative cues predicted the relevant location of probe stimuli that had been present in the array with only 80% validity. Valid cues pointed to the correct location of the probe in the previous array, whereas invalid cues pointed to an incorrect location that was previously occupied by another stimulus in the array. There were three types of cues: valid, invalid and neutral. To ensure sufficient trials in the invalid-cue condition, the number of SOAs between array and cue presentation was reduced to three – 150ms, 1200ms and 9600ms. These still spanned both IM and VSTM retention intervals.

There were 432 trials (360 informative and 72 neutral, 144 total at each SOA). Of the informative trials, 144 were valid (probe was in the array, at the cued location), 36 were invalid (probe was in the array, at an un-cued location), and 180 contained a probe stimulus that was absent from the array. Of the neutral trials, 36 contained a probe that was in the array and 36 contained a probe that was not in the array. There were 12 blocks of trials in the experiment, plus an additional practice block at the beginning.

3.3. Results

Accuracy data were converted into d’ scores, as in Experiment 1. These scores and the RT data were each analysed using an ANOVA, testing for behavioural benefits and costs of valid and invalid cues relative to neutral cues at the different SOAs. Since our validity manipulation only works for target-present trials, only these were included in our analyses (a target-absent trial is necessarily neither validly nor invalidly cued).

3.3.1. Accuracy

The d’ scores were compared using a two-way ANOVA (these scores can be seen in Figure 3). There was a main effect of validity [F(2,22)=4.647, p=0.021]. The d’ scores were higher with valid [1.97 ± 0.22 (Std.E.)] than with invalid [1.56 ± 0.20 (Std.E.)] [F(1,11)=5.984, p=0.032] or neutral [1.49 ± 0.32 (Std.E.)] retro-cues [F(1,11)=9.856, p=0.009]. The d’ scores for invalid and neutral trials did not differ significantly [F(1,11)=0.135, p=0.720]. There was a tendency towards an effect of SOA [F(1.37,15.12)=2.934, p=0.098]; the pattern showed a linear trend of decreasing d’ scores with increasing SOA, but this did not reach significance [F(1,11)=3.380, p=0.093].

The interaction between validity and SOA also did not reach significance [F(4,44)=2.150, p=0.090], although cue validity did significantly interact with the quadratic effect of SOA [F(1,11)=9.187, p=0.011]. To isolate the benefits and costs of valid and invalid cues relative to neutral retro-cues, we produced difference scores and compared these across the various SOAs. The relative benefit of valid retro-cues was significantly modulated by SOA [F(2,22)=5.158, p=0.015]. Follow-up t-tests revealed that the benefit at SOA 1200ms was greater than at SOA 9600ms [t(11)=4.216, p=0.001], but not significantly greater than at 150ms [t(11)=1.505, p=0.161]. There was no significant difference between the benefit at SOA 150ms and SOA 9600ms [t(11)=1.411, p=0.186].

3.3.2. Reaction Times

Only RTs taken from correct trials were used in this analysis. The RT data were submitted to a repeated-measures ANOVA, as in the d’ analysis. The ANOVA revealed a main effect of cue-type [F(2,22)=21.914, p<0.001]. Post-hoc contrasts revealed that all of the cue conditions differed significantly from one another. RTs were faster in valid trials [703ms ± 70ms (Std.E.)] than in neutral trials [834ms ± 55ms (Std.E.)] [F(1,11)=18.1, p=0.001] and invalid trials [893ms ± 68ms (Std.E.)] [F(1,11)=32.3, p<0.001]. RTs in invalid trials were slower than in neutral trials [F(1,11)=6. 534, p=0.030]. The pattern demonstrated both benefits of valid spatial cues and costs of invalid spatial cues compared to neutral trials throughout the intervals tested. There was also a main effect of SOA [F(2,22)=6.534, p=0.006]. Post-hoc contrasts indicated slower performance at the 9600ms SOA [877ms ± 70ms (Std.E.)] than at the 1200ms SOA [761ms ± 51ms (Std.E.)] [F(1,11)=23.3, p=0.001], and marginally slower than at the 150ms SOA [792m ± 58ms (Std.E.)] [F(1,11)=4.3, p=0.06]. Performance at the two shorter SOAs (150ms and 1200ms) differed, though this was only marginally significantly [F(1,11)=3.6, p=0.08]. There was no significant interaction between cue and SOA [F(4,44)=0.353, p=0.840].

3.4. Discussion

Whilst some previous studies have incorporated invalid trials (Astle et al. 2010; Griffin & Nobre, 2003; Matsukura et al. 2007), to our knowledge this is the first to explore the extent to which these costs are modulated by cue SOA. Subjects were faster to find probe-matching coloured shapes in memory when they have been validly cued, relative to when a neutral cue was used. They were also slower to find those shapes when they had been invalidly cued, also relative to a neutral cue baseline. This is the case across all SOAs. However, the d’ scores tell a subtly different story: manipulating both SOA and cue validity appears to have had some unintended consequences. Because the array size was always within ‘capacity’ (=< 4 items), subjects could strategise their performance; by introducing invalid cues, one possibility is that participants attempt to maintain all items in case the cue transpires to be invalid. The d’ scores would seem to support this; there is no d’ cost of invalid relative to neutral cues, implying that subjects were not using the retro-cue to discard un-wanted items. If participants had already maintained the contents of VSTM for 9600ms, it would be potentially wasteful for them to select a single item for the last 500-1000ms when there is a possibility that they would have selected the wrong item – thereby reducing the effect of the retro-cue at the longer SOA. Indeed, not only was there no invalidity cost at SOA 9600, there was no validity benefit either – in marked contrast to Experiment 1.

The lack of accuracy costs on invalid trials relative to neutral trials in Experiment 2 would seem to contrast with those of Matsukura et al. (2007), who demonstrated that invalid cues had a detrimental effect on accuracy measures, both with a double-cue and single-cue paradigm. One possibility is that subjects use the cue in a strategic way, only discarding items if it is essential, and certainly not doing so if they have already maintained all the array items for >9 seconds. That is, subjects might not always use a retro-cue to ‘protect’ the cued VSTM item at the expense of the other items.

To summarise briefly: The RT data suggest that subjects use the retro-cue to direct their memory search to the most likely probe location, thereby speeding RTs on valid trials and slowing RTs on invalid trials. However, there is relatively little effect of cue validity on the accuracy data, and certainly there is no accuracy cost of invalid cueing. This implies that subjects did not use the retro-cue to discard un-cued items, and thus reduce the load (as in Matsukura et al. 2007), but rather to guide their memory search. Given that the array size is always within VSTM capacity [Cowan, 2001], when the cue is invalid on some trials it is safest for participants to encode *all* items into VSTM – thereby reducing the relative benefit of a valid retro-cue, and eliminating any cost of invalid retro-cueing. Were this to be the case, we might expect the relative benefits and costs of retro-cueing to change as a function of load. When the array-size exceeds capacity, participants may have no choice but to use the cue to discard un-cued items. The following experiment tested this possibility.

4.1: Experiment 3: Orienting attention to locations within large arrays in visual short-term memory

A traditionally held view is that VSTM, as assessed using a change detection paradigm, has a capacity of around four items [Cowan, 2001; Luck & Vogel, 1997; Vogel & Machizawa, 2004]. The size of the memory array is varied, and the proportion of non-changes incorrectly detected (‘false alarms’) is subtracted from the proportion of changes detected (‘hits’), and then multiplied by the array size to produce an estimate of the number of items stored per array-size, or K [Cowan, 2001]. At around four items, participants’ K typically asymptotes, implying that even though the arrays are getting progressively larger, the extra items are not being stored in VSTM. In some cases the K estimate starts to decrease at around 4 items, implying that not only are the extra items not being stored, but that they are actively interfering with the storage (or retrieval) of the other items.

In one previous study using spatial retro-cues, the authors argued that the true capacity of VSTM could be higher (even above ten items) prior to the allocation of spatial attention, but that this fragile representation is overwritten at the probe onset [Sligte et al. 2008]. When a retro-cue is used participants are able to insulate the cued item from this overwriting (see also Makovski et al. 2008). Accordingly, they demonstrated that when the retro-cue is always valid, it could boost VSTM capacity to around 15 items. In short, they suggest that capacity is much higher than four items, provided that you can insulate those items from over-writing with spatial attention (as directed with a retro-cue). However, it is likely that the effect of the spatial retro-cues will be modulated by load. When load is low, or at least ‘within capacity’, subjects may use the retro-cue to prioritise memory search rather than discard items – as in Experiment 2. However, when load is high, or at least ‘exceeds capacity’, subjects may use the cue to discard un-cued items from VSTM (as in Matsukura et al. 2007), thereby reducing the memory load to a single item. Experiment 3 explores this possibility.

In Experiment 3 we sought to explore two things: i) we wanted to confirm previous findings that VSTM capacity is considerably higher, as measured using ‘supra-capacity’ arrays, following valid cues [Sligte et al. 2008]; ii) we sought to extend these findings by comparing both benefits and costs of spatial retro-cueing across different loads, to see whether the function of the cue changed depending upon the array size. Experiment 3 used a VSTM task more typical of those experiments explicitly focused on the limits of VSTM capacity, with subjects having to detect changes at a probed location, rather than to detect the presence of a probe at any location (the reasons for this change are explained in the Materials and Methods section).

4.2. Materials and methods

The main new manipulation was the introduction of large arrays containing 8 items, which are thought to exceed VSTM capacity [Cowan, 2001; Luck & Vogel, 1997; Vogel & Machizawa, 2004].

4.2.1. Participants

Ten participants (age range: 19 – 32 years, 8 females) took part in the experiment as volunteers.

4.2.2. Stimuli and Task

The task is shown in Figure 4. In this experiment, the stimulus array was composed of 2, 4 or 8 differently coloured crosses (33.3% probability). As in previous experiments, each trial began with a fixation point (400-700ms), followed by the array of crosses (100ms). Then, after a random interval between 1500 and 2500ms, a cue was presented on the screen (100ms) and then after 500-1000ms the probe stimulus was displayed (100ms).

To make our paradigm more comparable to those typically used to estimate VSTM capacity, the probe appeared peripherally, at one of the previously occupied locations and participants were required to perform a change-detection task. Participants' task was therefore to decide whether the probe colour was the same as, or different from, that of the array shape previously presented at that location. They were instructed to decide (yes or no) whether the colour of this probe matched the colour of the cross at that same location in the preceding array. In "no" trials, the probe item was always a colour that had been present in the array, but that had appeared in a different location.

The change in paradigm from memory search (Experiments 1 and 2) to change-detection carried several advantages for addressing the specific experimental questions. Any behavioural differences between valid, invalid and neutral trials cannot be affected by there being differential response criteria for different array locations, since decisions are probed at each location separately [Griffin & Nobre, 2003]. In addition, the task ensures that participants relied on the memory about the item at a particular location, rather than on a general sense of familiarity or novelty about a particular colour. One additional, pragmatic reason is that, as load increases, the stimulus set size also increases in an unwieldy way. To maintain a memory-search task, we would need sixteen differently coloured items, making it difficult to differentiate the item colours because they become so similar. With the paradigm used for Experiment 3, only eight items are needed, greatly reducing this problem. One potential difficulty with the change-detection paradigm is that subjects may make intrusion errors in this type of paradigm – i.e. that whilst item information is maintained, locations might become confusable, especially when changes occur at adjacent locations (see Chow, 1986, for a description).

There were three main trial types, valid, invalid and neutral. Informative cues in this task predicted (80% validity) the location that would be probed. The cue consisted of two overlapping white squares whose corners each pointed to one of 8 possible locations in the array. Valid and invalid informative cues (80% validity) consisted of one corner of the cue being filled in, pointing to a location that had been occupied in the preceding array. On valid trials, the cue correctly predicted the location of the array item which would be probed. On invalid trials, the cue incorrectly predicted the location of the array item which would be probed. Both colour-match and colour-non-match trials could be validly or invalidly cued. In neutral trials, all corners of the cue shape were filled in and the cue gave no spatial information. For each array load (2, 4, 8 items) and cue type (valid, invalid, neutral) the probability of correct colour match and non-match was equal (50%).

There were 504 trials in total (336 valid, 84 invalid, 84 neutral). Of the valid trials there were 56 trials in each response and load condition (match-load 2, match-load 4, match-load 8, non-match-load 2, non-match-load 4, non-match-load 8). In invalid and neutral trials, there were 14 trials in each response and load condition (match-load 2, match-load4, match-load 8, non-match-load 2, non-match-load 4, non-match-load 8). There were 14 blocks of trials in the experiment, plus 1 additional practice block at the beginning. As in previous experiments, no feedback was given during the experiment.

4.3. Results

The main aim of this experiment was to investigate the effect of retro-cues across the different VSTM loads. RT analyses focused on trials when one of the items had changed colour only, though the K estimate and d’ calculations also include no-change trials. Accuracy and RTs were analysed by repeated-measures analyses of variance (ANOVAs) testing the factors of cue (valid, invalid, neutral) and load (2, 4 or 8 items).

4.3.1. Accuracy

We produced d’ scores and submitted these to a two-way ANOVA, with the factors of load and validity. There was a significant effect of load [F(2,18)=37.799. p<0.001], with there being a significant linear trend of decreasing d’ scores with increasing load [F(1,9)=78.767, p<0.001]. D’ at load-2 [3.75 ± 0.45 (Std.E.)] was significantly higher than that at load-4 [2.48 ± 0.33 (Std.E.)] [F(1,9)=12.247, p=0.007] and than that at load-8 [0.96 ± 0.28 (Std.E.)] [F(1,9)=104.444, p<0.001]. D’ scores were also higher for load-4 than for load-8 [F(1,9)=31.351, p<0.001]. However, there was also a main effect of validity on d’ scores [F(1.31,11.75)=4.503, p=0.048], with significantly higher scores on valid than on invalid trials [F(1,9)=7.933, p=0.020], but no significant difference between neutral and invalid trials [F(1,9)=0.363, p=0.561]. There was a significant decrease in d’ scores from valid to neutral [F(1,9)=13.287, p=0.005]. Importantly, there was also a significant interaction between validity and load [F(4,36)=3.100, p=0.027]. There was no significant effect of cue-type on Load-2 trials [F(2,18)=1.123, p=0.347], but there was on both Load-4 [F(2,18)=5.150, p=0.017] and on Load-8 [F(2,18)=13.120, p<0.001] trials. At Load-4 the effect stemmed from a significant benefit for valid retro-cueing, relative to either the invalid [t(9)=2.377, p=0.041] or neutral [t(9)=2.739, p=0.023] condition. There was no significant difference between the invalid and neutral conditions [t(9)=0.381, p=0.712]. At Load-8 the effect stemmed instead from a significant cost for invalid retro-cueing, relative to both the valid [t(9)=6.019, p<0.001] and neutral conditions [t(9)=4.045, p=0.003]. At Load-8 there was no significant difference between valid and neutral conditions [t(9)=0.832, p=0.427].

4.3.2. Capacity estimates

We also used our accuracy data to produce K estimates. These were calculated by subtracting the proportion of false-alarms from the proportion of correct hits, for each load, and then multiplying this by the set size [Cowan, 2001]. The peak in K across the various loads is usually taken as an estimate of that participant's capacity [Fukuda & Vogel, 2009; Vogel & Machizawa, 2004], but in this case we included load as a factor. These data can be seen in Figure 5. The statistics are reported using the Greenhouse-Geisser correction for non-sphericity. There was a significant main effect of load [F(1.25,11.27)=4.567, p=0.049], though this was to be expected given that load was included in the calculation of K. More interestingly there was a significant main effect of cue-type [F(1.38,14.43)=7.423, p=0.012], with K being higher on valid than on invalid trials [F(1,9)=18.398, p<0.001] and higher on valid than on neutral trials [F(1,9)=5.519, p=0.043]. Invalid trials did not incur a significant cost relative to neutral trials [F(1,9)=2.708, p=0.134]. Importantly, these two factors interacted significantly [F(2.30,20.69)=7.371, p=0.003]. This was because there was no effect of cue-type at load-2 [F(1.32,11.84)=2.059, p=0.177], but there was at load-4 [F(1.74,15.66)=5.845, p=0.015] and at load-8 [F(1.72,15.51)=7.961, p=0.005]. At load-4, valid retro-cues produced a higher K estimate than on either the neutral condition [t(9)=3.471, p=0.007] or the invalid retro-cue condition [t(9)=2.866, p=0.019]. At load-4 neutral and invalid cues did not differ significantly [t(9)=0.727, p=0.486]. At load-8 the significant effect of cue-type arose because invalid retro-cues produced a reduced K estimate relative to valid retro-cues [t(9)=3.775, p=0.004] and to neutral cues [t(9)=2.829, p=0.020]. At load-8 neutral and valid retro-cues did not differ significantly [t(9)=0.276, p=0.789].

In summary: the costs and benefits of invalid and valid retro-cues were modulated by load. At Load-4 participants derived a relative benefit from the valid retro-cues, but no cost from the invalid retro-cues. At Load-8 valid retro-cues did not provide a benefit relative to the neutral retro-cues, but invalid retro-cues did elicit a significant cost. When the load is within VSTM capacity (load-4) valid spatial retro-cues provide a benefit relative to both invalid and neutral cues. By contrast when VSTM capacity is 'exceeded' there is a significant cost to invalid retro-cueing relative to both the valid and neutral conditions. This pattern of effects is apparent both in the d' and K scores.

4.3.3. Reaction Times

Only RTs from correct trials were used in this analysis. The pattern of RTs is shown in Figure 5. The ANOVA showed a main effect of cue [F(2,18)=106.210, p<0.001]. Post-hoc contrasts revealed that responses were significantly faster on valid trials [608ms ± 45ms (Std.E.)] compared to neutral trials [806ms ± 53ms (Std.E.)] [F(1,9)=184.614, p<0.001] and invalid trials [967ms ± 63ms (Std.E.)] [F(1,9)=138.988, p<0.001]. Furthermore, RTs were significantly slower on invalid trials compared to neutral trials [F(1,9)=37.945, p<0.001]. Load also exerted a main effect [F(2,18)=37.373, p<0.001], with there being a significant linear trend for increasing RTs with increasing load [F(1,9)=73.119, p<0.001]. Post-hoc analyses showed that RTs slowed significantly for each load increment. Responses in load-2 trials [695ms ± 56ms (Std.E.)] were significantly faster than in load-4 trials (794ms ± 76ms (Std.E.)) [F(1,9)=27.414, p=0.001] and load-8 trials [895ms ± 67ms (Std.E.))

[F(1,9)=73.119, p<0.001]. Load-4 trials also had significantly faster RTs than load-8 trials [F(1,9)=14.463, p=0.004]. There was no significant interaction between validity and load [F(4,36)=1.098, p=0.373].

4.4. Discussion

The results show that spatial orienting to items within VSTM representations modulated retrieval processes. As with cue SOA in the previous experiments, the relative RT benefits and costs of valid and invalid cueing were significant across all levels of load. However, in contrast with previous studies [e.g. Sligte et al. 2008], we only observed significant d' benefits when the array size was within capacity; when the array size was beyond capacity (8 items), valid cueing inferred no benefit relative to neutral trials. Conversely, only when capacity was exceeded did we observe a substantial *cost* of invalid cueing. This pattern of effects was present both in both the d' and K measures.

A recent study demonstrated that the capacity prior to attention being committed is much higher than the traditional four items. Sligte et al [2008] compared performance on retro-cue and post-cue conditions and demonstrated that the capacity of VSTM is at least double the traditional VSTM capacity, but that the apparent capacity drops to four items only on the presentation of the probe array. However, the results of Experiment 3 seem inconsistent with this view; valid retro-cues should have provided a benefit even when capacity was 'exceeded', whereas the benefit is limited to within-capacity loads. Not only were capacity estimates never boosted beyond 4 items, when the array size was 8 items the valid retro-cue conferred no significant benefit, relative to neutral trials.

An additional contribution of Experiment 3 is that it explores the relative *costs* of this retro-cue-driven orienting. The pattern of results suggests that when load exceeds capacity, subjects use the cue to discard those un-cued items from VSTM, having a catastrophic effect on accuracy when the cue is invalid. Indeed subjects were worse than were they to have attempted to maintain all the array items (neutral trials). The results of Experiment 3 would therefore seem to support and extend the findings of Matsukura et al. (2007): the use of the retro-cue differs depending upon whether VSTM capacity is exceeded; only when it is exceeded will subjects resort to using the retro-cue to discard items.

At load 2 there appears to be little effect of cue validity. This is unlikely to be because cueing benefits cannot be observed at such a low set-size (see Nobre et al. 2008), but because participants opt to use the cue only when the benefits of doing so outweigh the potential costs. If subjects can maintain the items in VSTM, then there may be little incentive to use the cue, given that it might transpire to be invalid. Again, we suggest that subjects use the retro-cue strategically; if there were no potential cost to using the retro-cue we would expect a cueing benefit, even at load 2 (as in Nobre et al. 2008).

Matsukura et al. observed the relative cost of a single invalid cue, and an initial invalid cue in the double-cue paradigm, at both load 4 and load 6. This would seem to contrast with our interaction between VSTM load and cue validity. Why would subjects use the cue differently within and beyond VSTM capacity in our Experiment 3, and yet not do so in

Matsukura et al.’s Experiment 3? We suspect that to get this pattern of costs across different loads, subject need to experience all loads. Matsukura et al.’s comparison of load 4 and load 6 was between-subjects, whereas our comparison of load 4 and load 8 is within-subjects. Subjects may only demonstrate a differential strategy of cue use if they themselves experience the different memory loads. Furthermore, Matsukura et al. did not include any neutral trials, making it difficult to ascertain whether they are observing a validity benefit or an invalidity cost.

5.1. Experiment 4: Orienting attention within large arrays in IM and VSTM

Across the preceding three experiments we have explored two potential constraints upon our ability to orient attention within mental representations. The first was the temporal constraint of whether the representation was held in IM or VSTM, and for how long the array had been held in VSTM. The second was the constraint of load. Experiment 3 showed a very interesting pattern of accuracy costs and benefits across load, with relative benefits, but no costs, when cueing at capacity, and the reverse pattern when load exceeds capacity. Experiment 4 explores whether or not this relationship between load and retro-cueing benefits / costs interacts with cue SOA, by investigating whether similar modulations across load are observed within IM time spans.

Previous research has demonstrated that the capacity of IM and VSTM differ greatly [Averback & Coriell, 1961; Sperling, 1960]. This was supported by a recent study showing that a valid retro-cue could boost VSTM capacity estimates to around 15 items, but that when operating within IM, K could reach around 20 items. When the stimuli leave an afterimage, the apparent capacity of IM could be even higher [Sligte et al. 2008]. With this in mind, our first aim was to replicate this effect and test whether capacity measures could be boosted most when the retro-cue operates on an IM, relative to a VSTM, representation. Secondly we aimed to test whether the pattern of costs across load seen in Experiment 3 also occurred when the retro-cue operated upon an IM representation.

5.2. Materials and methods

Experiment 4 used high-load arrays (8 items) in addition to the variation of the interval between the array and the cue, including retention intervals spanning IM and VSTM intervals. Unless stated otherwise, the materials and methods are identical to those in Experiment 3.

5.2.1. Participants

Ten participants (age range: 19 – 32 years; 8 female) took part in the experiment as volunteers.

5.2.2. Stimuli and Task

The basic task is shown in Figure 4. The main difference in this experiment compared to Experiment 3 was that the interval between the appearance of the array and the cue was either 150ms (IM) or 1500ms (VSTM). In addition, only two array loads (4 or 8 items, 50% probability) were used. Cues could be valid, invalid or neutral.

There were 480 trials in total (320 valid, 80 invalid, 80 neutral). For valid cues, there were 40 trials in each experimental cell (match-load4-SOA150ms, non-match-load4-SOA1500ms, non-match-load4-SOA150ms, non-match-load4-SOA1500ms, match-load8-SOA150ms, non-match-load8-SOA1500ms, non-match-load8-SOA150ms, non-match-load8-SOA1500ms). For invalid and neutral cues, there were 10 trials in each cell.

5.3. Results

As in Experiment 3, the RT analysis focused on the trials upon which the probe item colour had changed (i.e. non-match trials), but the K estimate and d’ calculations used both correct hits and false alarms. All of the effects were assessed by ANOVAs testing the factors of cue (valid, invalid, neutral), load (4, 8 items), and SOA (150, 1500ms).

5.3.1. Accuracy

The accuracy scores were converted into d’ scores. We entered these scores to a three-way ANOVA, with the within-subjects factors of SOA, Load and Validity. There was a significant main effect of load, with scores being higher for load-4 [2.02 ± 0.40 (Std.E.)] than for load-8 [0.88 ± 0.38 (Std.E.)] [F(1,9)=56.536, p<0.001]. There was a significant main effect of validity [F(2,18)=57.518, p<0.001], with scores being higher for valid [2.26 ± 0.36 (Std.E.)] than neutral [1.45 ± 0.41 (Std.E.)] [F(1,9)=52.547, p<0.001] and for valid than invalid [0.65 ± 0.37 (Std.E.)] [F(1,9)=109.299, p<0.001] trials. D’ scores on invalid trials were also significantly lower than on neutral trials [F(1,9)=20.230, p<0.001]. The only significant interaction was between SOA and validity [F(2,18)=7.274, p=0.005]. This resulted from a significant validity effect at SOA 150ms [F(2,18)=37.130, p<0.001] that stemmed both from a benefit of valid relative to neutral retro-cues [F(1,9)=50.799, p<0.001], and of a cost of invalid relative to neutral retro-cues [F(1,9)=11.241, p=0.008]. There was also a significant difference between valid and invalid conditions [F(1,9)=48.797, p<0.001]. At SOA 1500ms, the significant effect of validity [F(2,18)=29.264, p<0.001] was driven by only a small benefit of valid relative to neutral retro-cues [F(1,9)=5.356, p=0.046], but a substantial cost of invalid retro-cueing relative to the neutral condition [F(1,9)=17.716, p=0.002]. There was also a significant difference between valid and invalid conditions [F(1,9)=114.594, p<0.001].

5.3.2. Capacity measure

We also included the K estimates across load in a repeated-measures ANOVA, with the factors of SOA, validity/cue-type and load (these data can be seen in Figure 6). There was a significant main effect of cue-type [F(2,18)=56.342, p<0.001], with K on valid trials being higher than on neutral [F(1,9)=56.112, p<0.001] and invalid trials [F(1,9)=66.699, p<0.001]. There was no main effect of Load [F(1,9)=0.005, p=0.947] or SOA [F(1,9)=0.959, p=0.353], however there were a number of 2-way interactions: there was a significant interaction between SOA and cue-type [F(2,18)=6.480, p=0.008], with the cues having a greater influence on K at the IM interval [F(2,11)=40.211, p<0.001] than at the VSTM interval [F(2,18)=21.255, p<0.001]. At the IM delay/interval the effect of retro-cues was mainly carried by a relative benefit for valid relative to neutral [F(1,9)=51.963, p<0.001] and invalid conditions [F(1,9)=48.535, p<0.001]. With the cost of invalid retro-cues, relative to the neutral condition, also being significant [F(1,9)=8.882, p=0.015]. At the VSTM delay, the effect of retro-cues was also carried by significant differences between valid retro-cues and the neutral condition [F(1,9)=10.920, p=0.001], and with the invalid condition [F(1,9)=28.248, p<0.001]. There was also a significant K cost for invalid retro-cues relative to the neutral baseline [F(1,9)=8.428, p=0.018]. There was a significant interaction was between load and cue-type [F(2,18)=6.353, p=0.008], because load had the greatest effect on valid trials [t(9)=−2.362, p=0.042], with there being no significant effect at of load on neutral [t(9)=0.839, p=0.423] or invalid trials [t(9)=1.902, p=0.090]. The 2-way interaction between SOA and load approached significance [F(1,9)=4.473, p=0.064], this was because whilst load did not have a significant effect on K estimates at either SOA, it had a positive effect on K in the IM interval [t(9)=−1.419, p=0.190], and, as we had seen in the previous experiment, it had a negative effect on K at the VSTM interval [t(9)=1.663, p=0.131]. There was no 3-way interaction between SOA, cue-type and load [F(2,18)=1.873, p=0.182].

In summary: retro-cueing modulated accuracy both in the IM and VSTM intervals. As in Sligte et al. [2008], the retro-cues had their greatest effect at the IM delay, although our capacity estimates were not boosted by nearly the same extent. This modulation was driven primarily by a relative benefit from valid spatial retro-cueing, especially when the array size exceeded VSTM capacity. The cue-type also exerted a significant effect at the VSTM delay.

5.3.3. Reaction Times

Only RTs from correct trials were included in this analysis. The RT results are shown in Figure 6. The ANOVA showed a main effect of cue [F(2,18)=44.656, p<0.001]. Post-hoc comparisons showed that RTs were significantly faster in valid [576ms ± 55ms (Std.E.)] compared to neutral [846ms ± 72ms (Std.E.)] [F(1,9)=48.549, p<0.001] or invalid trials [938ms ± 102ms (Std.E.)] [F(1,9)=116.697, p<0.001]; and that RTs were significantly slower in invalid trials compared to neutral trials [F(1,9)=5.970 p=0.037]. Overall, RTs in load-4 (749ms ± 74ms (Std.E.)] trials were faster than in load-8 trials [824ms ± 86ms (Std.E.)] [F(1,9)=12.968, p=0.006]. The main effect of SOA was not significant, and there were no significant interactions.

5.4. Discussion

The pattern of results at Load-8 in the VSTM condition replicated the pattern of results in the previous experiment: valid retro-cues appear to confer little or no advantage relative to neutral retro-cues, but invalid retro-cues confer a substantial cost; this can be seen most obviously in our d’ estimates. However, the effect at the IM SOA was somewhat different: valid retro-cues substantially boosted performance at load-8, as well as invalid cues eliciting a significant cost. As was the case in Experiment 3, when operating on VSTM representations, valid cues enabled participants to achieve a capacity estimate of a little over three and half items, whereas when those valid retro-cues operated on IM representations participants were able to achieve a mean capacity estimate of over five items.

Sperling [1960] and others suggested that the capacity of IM is greater than that of VSTM. Accordingly, we might expect participants to be able to make better use of a valid retro-cue when they can use it to operate on the larger number of items held in IM than when operating on the already restricted items held in VSTM. As was mentioned in Experiment 1, the difference between retro-cueing at VSTM and IM intervals is most apparent when VSTM capacity is exceed. if valid retro-cues enable participants to operate upon fragile item traces [Averback & Coriell, 1961; Sligte et al. 2008; Sperling, 1960], then the use of such a cue might yield the greatest benefit when it can operate with IM rather than VSTM. Presumably these fragile supra-capacity traces are highly prone to decay or interference [Averback & Coriell, 1961; Sperling, 1960], meaning that they would be more available for selection by an early relative to a late spatial cue. Using invalid cues in Experiment 4, also enabled us to test whether the costs are equivalent across these two delays. They appear to be similar: d' scores drop close to chance when there are eight items in the array and cue is invalid, implying that across both VSTM and IM, when the size of the array exceeds VSTM capacity, subjects use the retro-cue to discard un-cued items, making them unavailable for re-inspection at probe-onset. Of course, the reason for this drop could differ across these two intervals: at IM SOA, the un-cued items might be lost to decay (Sperling, 1960); at the VSTM SOA the un-cued items might be lost intentionally, because capacity is exceeded (as in Experiment 3 and Matsukura et al. 2007).

6. General discussion

These experiments provide robust and compelling evidence that spatial attention can be oriented within the domain of mental representations to enhance and modulate memory retrieval. The four experiments focused on two constraints upon mental representations. The first of which was a temporal constraint. We compared attention orienting within IM and VSTM durations (Experiments 1, 2 and 4), and extended the VSTM duration to 9600ms to explore the effect of VSTM decay on retrospective attention orienting (Experiments 1 and 2). The second constraint was that of load. We explored the orienting of attention to locations within either VSTM (Experiment 3 and 4) or IM (Experiment 1, 2 and 4) representations, when they were either within the supposed capacity limit of VSTM (loads of 2 to 4 items) or beyond the capacity limit of VSTM (load 8). The final experiment brought these two constraints together, exploring any effects load on the orienting of attention within IM and VSTM representations. Our aim was to test whether these factors might influence the mechanisms through which retro-cues improve performance.

There are a number of important results; i) when the informative spatial retro-cue was 100% valid, and the set size was within VSTM capacity, it enhanced retrieval over a wide range of SOAs, biasing representations in both IM and VSTM, even at cue SOAs of ~10 seconds (Experiment 1); ii) when the number of to-be-remembered items did not exceed VSTM capacity, subjects did not use the retro-cue to discard un-cued items – even when cued to the wrong item subjects' retrieval is slower but no less accurate (Experiments 2, 3 and 4); iii) when the number of items exceeded VSTM capacity (8 items) subjects used the retro-cue to discard un-cued items – when cued to the wrong item, accuracy measures dropped to little better than chance levels (Experiments 3 and 4); iv) this pattern of benefits and costs, within and beyond capacity, did not hold when using the retro-cue to operate on an IM representation (Experiment 4). When operating within IM valid retro-cues confer an accuracy benefit both within and beyond 'capacity', indeed the greatest benefit is seen with 'supra-VSTM-capacity' arrays.

It has recently been shown that a valid retro-cue can provide a K benefit relative to a post-cue even at arrays of 32 items [Sligte et al. 2008]. Given this, it is surprising that whilst participants were *faster* to retrieve the validly cued item, we did not observe a K or d' benefit in arrays of eight items (Experiment 3). Our data do not therefore support the view that there exists a pool of items, greater than the capacity of VSTM, which can be maintained using a retro-cue. As we mentioned earlier, we suspect that one reason for this is that subjects will not rely on the cue to the same extent when they suspect that it might be invalid. Thus, whilst our designs enable us to explore the relative cost of retro-cueing to those items un-cued, they may also preclude us from observing the large benefits seen by Sligte et al. (2008). A possible secondary cause of this apparent difference is the degree of prior training that subjects were given. Because task instructions were relatively simple, we gave subjects around 50 trials practise; Sligte et al. (2008) gave their subjects substantially more training (around three hours). Highly trained subjects, who rely entirely on the retro-cue, might be required in order to see these massive increases in K estimates by retro-cues. However, a more recent study from the same group does fit well with our findings: Sligte et al. 2010 used 100% valid retro-cues (as in Experiment 1), delivered at different SOAs, to probe the capacity of different short-term stores. As in Experiments 3 and 4, the array size was 8 items. Delivering the cue within an IM period (10 ms post array) yielded K estimate of a little over 6 items, which dropped to a little over 4 items when the cue was delivered within a VSTM period (1000 ms post array), and dropped further to a little over 2 items in a non-retro-cued version (in this control condition the cue was delivered after the change had occurred). These K estimates are similar to those that we observed in Experiment 4: when operating on 8 items in IM (cue SOA 150 ms), retro-cues yielded a K estimate of a little over 5 items, which dropped to a little over 3 items when the cue was delivered in a VSTM period (1500 ms post array), and dropped further to around 2 items when no retro-cue was used (the neutral conditions). Our K estimates are a little lower than those observed by Sligte et al. (2010), but this is perhaps to be expected: Sligte et al.'s (2010) cues were delivered slightly earlier, and they used real-life objects rather than easily-confusable coloured crosses.

Our results extend previous findings in two respects: Firstly, we used neutral cues in order to establish the *relative benefit* of these different retro-cue SOAs. The retro-cue is most effective at boosting capacity estimates when it operates upon IM, consistent with the view that these items are most easily lost without such cues (Sperling, 1960; Averback and Corriel, 1961). The retro-cue provides little *benefit* in terms of boosting capacity when it operates upon a VSTM representation. Secondly, our use of invalid cues demonstrates that retro-cues operating on 'supra-capacity' arrays are effective because, at least partially, they enable subjects to discard un-cued items, thereby reducing the load. This is true across all SOAs.

As with previous studies it is difficult to ascertain how items are lost from VSTM, and we certainly make no claim here as to how this arises. However, to our knowledge, for the first time, the current study demonstrates that the retro-cue is not always used in the same way. When VSTM capacity is not exceeded, subjects use the retro-cue to prioritise where they initiate their memory search, with retrieval being slowed but not less accurate following an invalid retro-cue (see also Nobre et al. 2008). It is only when VSTM capacity is exceeded that subjects use the cue to reduce the number of stored items. Although it remains to be seen how subjects do this, and whether the removal of items is achieved through the same means in IM and VSTM. One possibility is that un-cued items in IM are not insulated from decay, which, given the fragile nature of these representations, will mean that they are lost by probe onset [Sperling et al. 1960]; by contrast when these items are stored in VSTM, subjects may actively remove the previously applied insulation, or actively suppress the item representation.

A final important issue to consider, which may differentiate this study from some previous examples in the literature, is whether subjects emphasize speed or accuracy. In our experiments we did not emphasize one over the other, with us looking for cueing effects in both speed and accuracy. Other examples (e.g. Matsukura et al. 2007) have explicitly emphasized accuracy over speed to subjects. This difference may well affect the mechanisms at play. We are confident that our effects of load, validity of SOA in d' or K are not stemming simply from differential speed-accuracy trade-offs across these different factors; in no case do the RT effects reflect the inverse of the accuracy effects.

Acknowledgments

D.E.A. was supported by a Postdoctoral Fellowship from the British Academy and by project grants from the John Fell Fund (Oxford University Press) and the Royal Society. A.C.N. was supported by a 21st Century Research Award from the James S McDonnell Foundation (Bridging Brain, Mind and Behaviour) and by a project grant from the Wellcome Trust. We would like to thank Morag Brothwell for her help in collecting preliminary data that led to these experiments. We would also like to thank Bo-Cheng Kuo, Ilja Sligte and three anonymous reviewers for their comments on earlier drafts.

References

Astle DE, Nobre AC, Scerif G. Attentional control constrains visual short-term memory: Insights from developmental and individual differences. Q J Exp Psychol. 2010 (iFirst).

Astle DE, Scerif G, Kuo B-C, Nobre AC. Spatial Selection of Features within Perceived and Remembered Objects. Front Hum Neurosci. 2009; 3:6. [PubMed: 19434243]

Averback E, Coriell AS. Short-term memory in vision. Bell Systems Technical Journal. 1961; 40:309–328.

Bays PM, Catalao RF, Husain M. The precision of visual working memory is set by allocation of a shared resource. J Vis. 2009; 9(10):7, 1–11. [PubMed: 19810788]

Cowan N. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav Brain Sci. 2001; 24(1):87–114. discussion 114-85. [PubMed: 11515286]

Chow SL. Iconic memory, location information, and partial report. J Exp Psychol: Hum Percept Perform. 1986; 12(4):455–465. [PubMed: 2946802]

Dick AO. Iconic memory and its relation to perceptual processing and other memory mechanisms. Percept Psychophys. 1974; 16:575–596.

Fukuda K, Vogel EK. Human variation in overriding attentional capture. J Neurosci. 2009; 29(27): 8726–33. [PubMed: 19587279]

Griffin IC, Nobre AC. Orienting attention to locations in internal representations. J Cogn Neurosci. 2003; 15(8):1176–94. [PubMed: 14709235]

Landman R, Spekreijse H, Lamme VA. Large capacity storage of integrated objects before change blindness. Vis Res. 2003; 43(2):149–64. [PubMed: 12536137]

Lee D, Chun MM. What are the units of visual short-term memory, objects or spatial locations? Percept Psychophys. 2001; 63(2):253–7. [PubMed: 11281100]

Lepsien J, Griffin IC, Devlin JT, Nobre AC. Directing spatial attention in mental representations: Interactions between attentional orienting and working-memory load. Neuroimage. 2005; 26(3): 733–43. [PubMed: 15955482]

Lepsien J, Nobre AC. Cognitive control of attention in the human brain: insights from orienting attention to mental representations. Brain Res. 2006; 1105(1):20–31. [PubMed: 16729979]

Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. Nature. 1997; 390(6657):279–81. [PubMed: 9384378]

Makovski T, Shim WM, Jiang YV. Interference from filled delays on visual change detection. J Vis. 2006; 6(12):1459–70. [PubMed: 17209748]

Makovski T, Sussman R, Jiang YV. Orienting attention in visual working memory reduces interference from memory probes. J Exp Psychol Learn Mem Cogn. 2008; 34(2):369–80. [PubMed: 18315412]

Matsukura M, Luck SJ, Vecera SP. Attention effects during visual short-term memory maintenance: protection or prioritization? Percept Psychophys. 2007; 69(8):1422–34. [PubMed: 18078232]

Neisser, U. Cognitive psychology. Appleton-Century-Crofts; New York: 1967.

Nobre AC, Griffin IC, Rao A. Spatial attention can bias search in visual short-term memory. Front Hum Neurosci. 2008; 1:4. doi: 10.3389/neuro.09.004.2007. [PubMed: 18958218]

O'Regan JK, Rensink RA, Clark JJ. Change-blindness as a result of 'mudsplashes'. Nature. 1999; 398(6722):34. [PubMed: 10078528]

Posner MI. Orienting of Attention. Q J Exp Psychol. 1980; 32:3–25. [PubMed: 7367577]

Postle BR, Awh E, Jonides J, Smith E, D'Eposito M. The where and how of attention-based rehearsal in spatial working memory. Cogn Brain Res. 2004; 20(2):194–205.

Rensink RA, O'Regan JK, Clark JJ. To see or not to see: The need for attention to perceive changes in scenes. Psychol Sci. 1997; 8:368–373.

Sligte IG, Scholte HS, Lamme VA. Are there multiple visual short-term memory stores? PLoS ONE. 2008; 3(2):e1699. [PubMed: 18301775]

Sligte IG, Scholte HS, Lamme VA. V4 activity predicts the strength of visual short-term memory representations. J Neurosci. 2009; 29:7432–7438. [PubMed: 19515911]

Sligte IG, Vandenbrouke ARE, Scholte HS, Lamme VAF. Detailed sensory memory, sloppy working memory. Front Psychol. 2010; 1:175. [PubMed: 21897823]

Sperling G. The information available in brief visual presentations. Psychol Monographs. 1960; 74:1–29.

Sternberg S. High-speed scanning in human memory. Science. 1966; 153(736):652–4. [PubMed: 5939936]

Vogel EK, Machizawa MG. Neural activity predicts individual differences in visual working memory capacity. Nature. 2004; 428(6984):748–51. [PubMed: 15085132]

Vogel EK, Woodman GF, Luck SJ. Storage of features, conjunctions and objects in visual working memory. J Exp Psychol Hum Percept Perform. 2001; 27(1):92–114. [PubMed: 11248943]

Wheeler ME, Treisman AM. Binding in short-term visual memory. J Exp Psychol Gen. 2002; 131(1): 48–64. [PubMed: 11900102]

Zhang W, Luck SJ. Discrete fixed-resolution representations in visual working memory. Nature. 2008; 453(7192):233–5. [PubMed: 18385672]

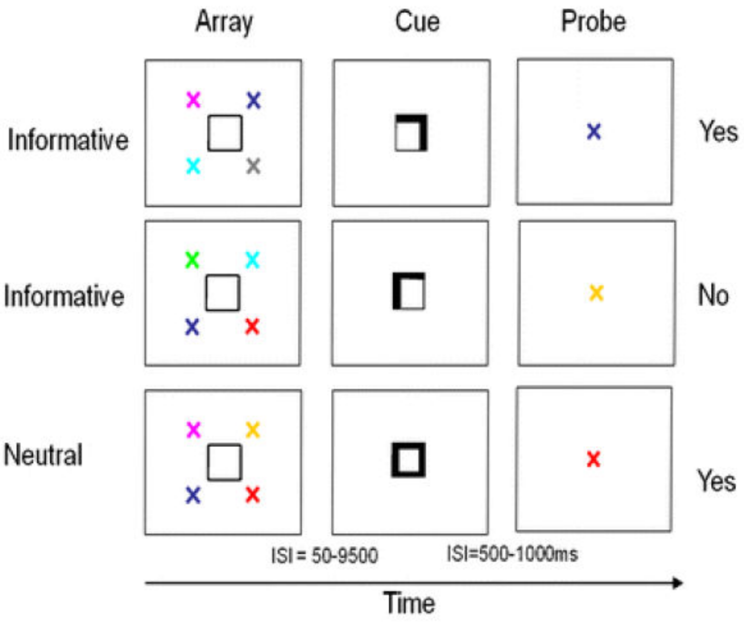

Figure 1.

A trial order schematic (Experiment 1), for two trials with informative spatial retro-cues and one neutral trial. The upper two trial schematics show probe present trials, the lower schematic shows a probe absent trial.

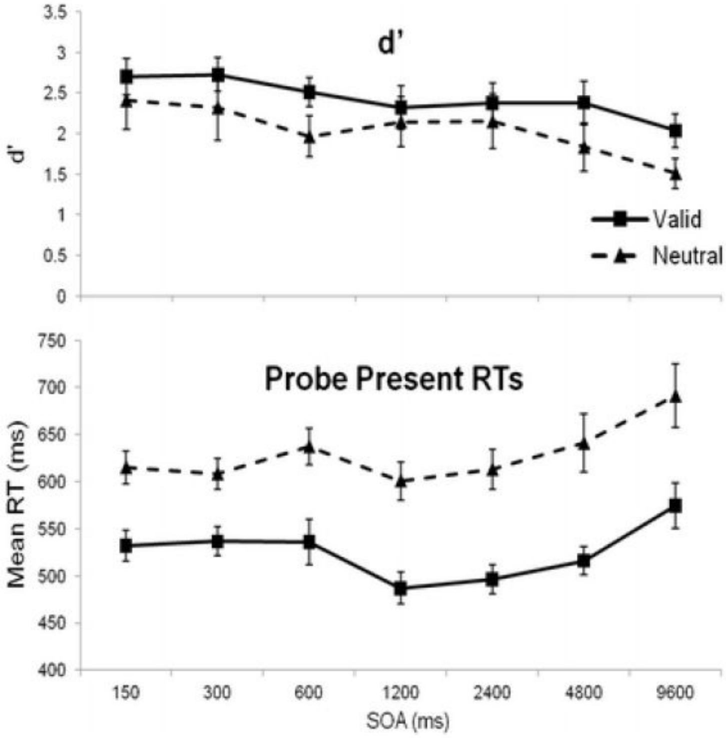

Figure 2.

Results from Experiment 1: the upper panel shows d prime scores for valid and neutral trials across the SOAs. The lower panel shows reaction times (RT) for the same comparison. In all cases the error bars show the standard error of the mean.

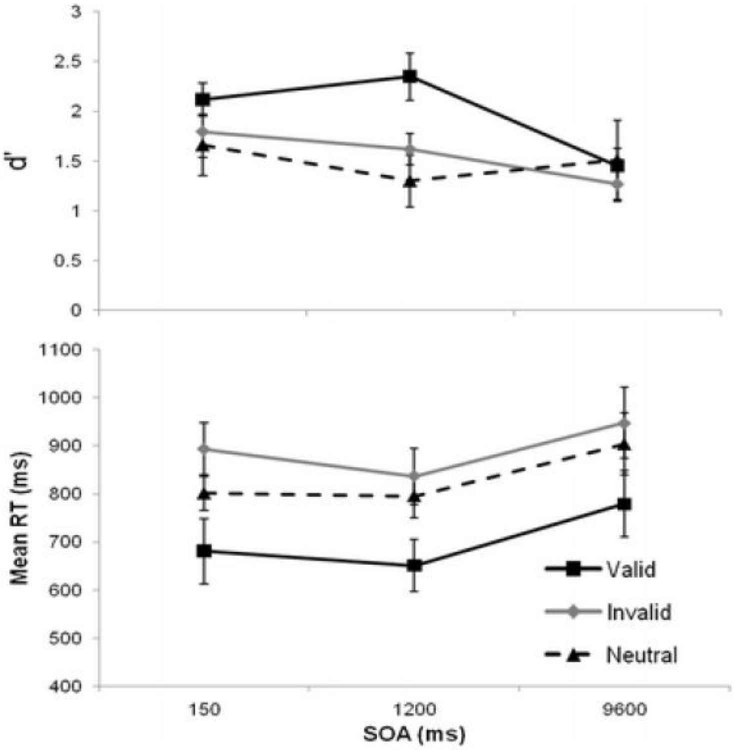

Figure 3.

Results from Experiment 2: the upper panel shows d prime scores for valid, neutral and invalid trials across the SOAs. The lower panel shows RT for the same comparison. In both cases the error bars show the standard error of the mean.

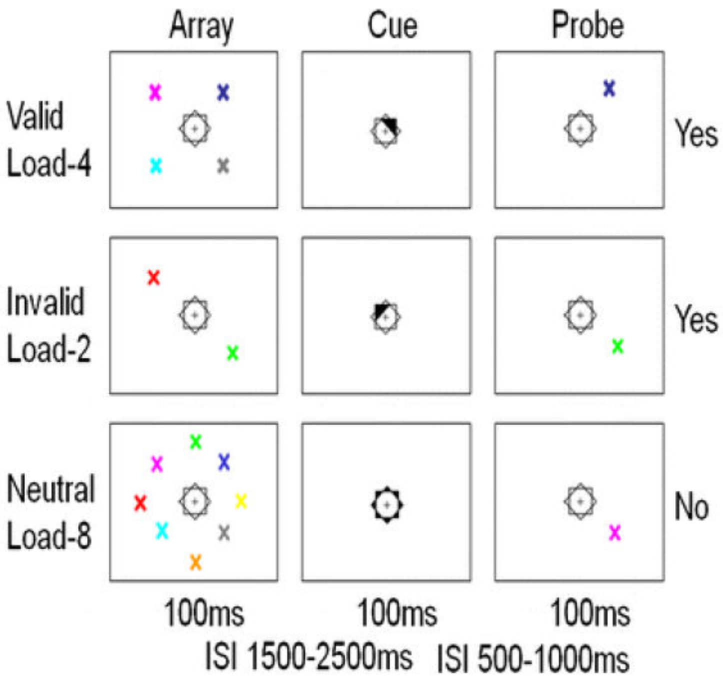

Figure 4.

A trial order schematic (Experiment 3), showing trials across three levels of load and three levels of cue-validity. The upper two schematics are probe present trials, the lower schematic is a probe absent trial.

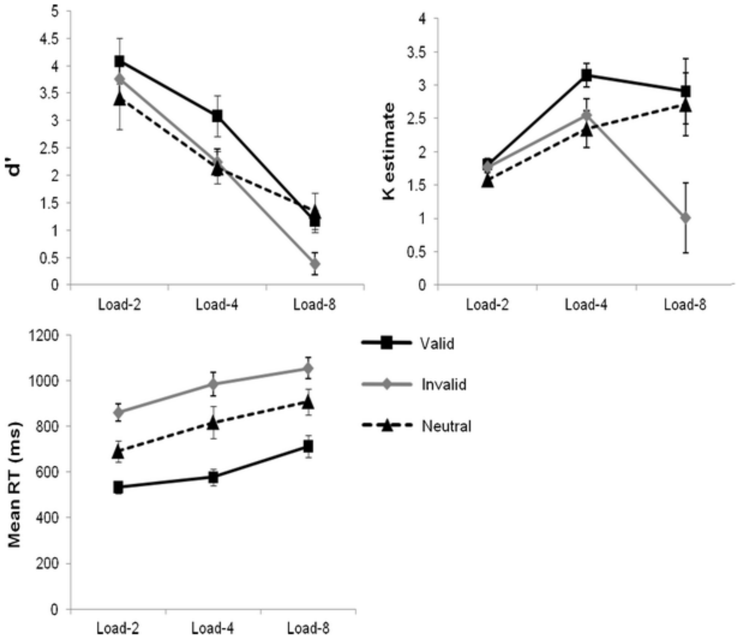

Figure 5.

Results from Experiment 3: the upper-left panel shows d prime across load, for the three levels of cue validity; the upper-right panel shows K estimates (proportion of hits minus proportion of false alarms, multiplied by load), for the same comparison; the lower panel shoes RT for the same comparison. In all cases the error bars show the standard error of the mean.

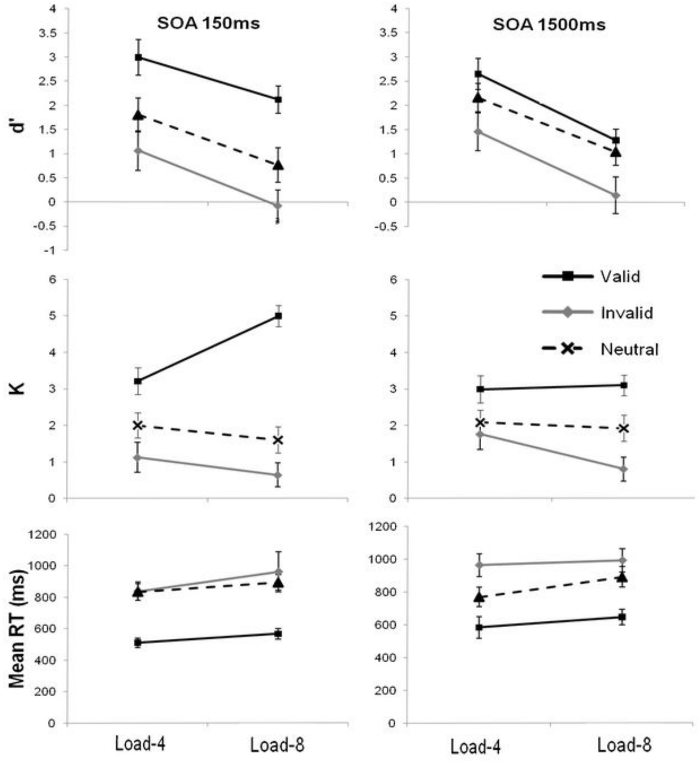

Figure 6.

Results from Experiment 4: the upper two panels show d prime across the two levels of load, for iconic memory (left) and VSTM (right) SOAs, with the three levels of cue validity; the middle two panels show the same comparison for K estimates (the proportion of hits minus proportion of false alarms, multiplied by load); the lower two panels show the same comparison for mean RT. In all cases the error bars show the standard error of the mean.