Memory & Cognition

Nothing to remember: Encoding and memory processes involved in representing empty locations

Journal:	*Memory & Cognition*
Manuscript ID	MC-ORIG-20-383.R2
Manuscript Type:	Resubmitted Manuscript
Date Submitted by the Author:	n/a

Complete List of Authors:	Csink, Viktoria; Birkbeck University of London, Psychological Sciences Gliga, Teodora; University of East Anglia, School of Psychology Mareschal, Denis; Birkbeck University of London, Psychological Sciences

Keywords:	empty locations, configural processing, change detection, pupillometry, visual-spatial short-term memory

Nothing to remember: Encoding and memory processes involved in representing empty locations

Csink, Viktoria [a], **Gliga**, Teodora [b], **Mareschal**, Denis [c]

[a] Birkbeck University of London, Department of Psychological Sciences, Malet St, Bloomsbury, London WC1E 7HX, United Kingdom, viktoria.csink@gmail.com, +44(0)7597769271

[b] University of East Anglia, School of Psychology, Norwich Research Park, Norwich NR4 7TJ, United Kingdom, T.Gliga@uea.ac.uk, +44 (0)1603 59 7967

[c] Birkbeck University of London, Department of Psychological Sciences, Malet St, Bloomsbury, London WC1E 7HX, United Kingdom, d.mareschal@bbk.ac.uk, +44 (0)20 7631 6582

Corresponding author: **Csink**, Viktoria

Keywords: empty locations, configural processing, change detection, pupillometry, visual-spatial short-term memory

Acknowledgements
The research was funded by the Centre for Brain and Cognitive Development, Birkbeck University of London. There are no relevant financial or non-financial competing interests to report.

Nothing to remember: Encoding and memory processes involved in representing empty locations

Abstract

Previous research has provided rich evidence that a set of visual objects can be encoded in isolation along with their exact coordinate positions as well as a global configuration that provides a network of interrelated spatial information. However, much less data is available on how unoccupied locations are encoded and maintained in memory. We tested this ability in adults using a novel paradigm that involved both empty and filled locations and required participants to monitor the addition or deletion of an item, which occurred 50% of the time. Crucially, a number of locations remained hidden to the participant, thus information on the absence of an item at a location could not be inferred from the presence of items elsewhere. We used eye-tracking to measure the proportion of target looking during encoding and the amount of pupil dilation during memory retention. Participants looked significantly longer at filled compared to empty targets and target looking during encoding only predicted accuracy in case of filled targets. Increased pupil dilation was observed in response to an increasing number of items while pupil diameter was unaffected by the number of empty locations. In addition, participants made significantly more errors in the conditions that involved the representation of an empty location. Our findings support the view that human adults encode exact coordinates of items in memory. In contrast, we suggest that empty locations are represented as a property of the global configuration of items and empty space, and not as independent units of information.

Keywords: empty locations, configural processing, change detection, pupillometry, visual-spatial short-term memory.

Introduction

A vast literature in vision science has investigated the attentional processes underlying the selection of objects in the visual scene, and the ability to encode and maintain a representation of discrete visual objects (for a review on object-based attention see Scholl, 2001). In contrast, very little is known about how empty space or discrete empty locations in the scene are represented, if at all.

Although the cognitive demand to represent items and their locations is undoubtedly more common than keeping track of unoccupied locations, a representation of the *lack* of a relevant item – or any item – at a location is both possible and useful when encoding the visual scene. Keeping an inventory of empty locations might be vital for encoding and retaining information on the number and size of available locations where items might later be placed. These processes are present during the visual search for a parking space while driving and maintaining the information on potential slots in the area until the most suitable space is chosen. Additionally, when searching for an item, retaining information on empty locations helps avoid these locations in future searches. For instance, when searching through boxes to find a specific item, each box that has already been opened needs to be tagged as 'empty' or not containing the item in order to differentiate them from the ones that have not been opened yet.

However, in the absence of any visual features that would individuate a spatiotemporally distinct object and anchor it to a specific location in space, it is difficult to understand how these representations are formed and what their specific content is.

The difficulty of representing empty locations or the absence of an item at a location is reflected in the infant cognitive literature that finds no evidence of surprise at 8 months of age when an item magically appears at a location that was shown to be empty only seconds earlier (Wynn & Chiang, 1998; Kaufman, Csibra & Johnson, 2003).

This finding is in stark contrast with the finding that infants as young as 3 months show a surprise response if an item magically *disappears* (Baillargeon, 1987; Baillargeon & DeVos, 1991; Kaufman, Csibra & Johnson, 2003; 2005). In the magical disappearance scenario, an object is first covered and then fails to appear when the cover is lifted (Wynn & Chiang, 1998; Kaufman, Csibra & Johnson, 2003). Infants might respond to this scenario with surprise because they maintained a lingering percept of the visual features of the object while it was in occlusion, therefore the outcome that presents no object violates this visual representation (Haith, 1998).

In contrast, in case of magical *appearance*, an empty target location is covered, followed by the appearance of the object at the empty location. In this scenario there is no discrete object at the target location to form and maintain a visual representation of, and therefore when the object is presented, there is no specific object representation in the infant's visual memory that the outcome contradicts. In other words, the new item at an implausible location does not contradict any prior representation if there was *nothing* to represent to begin with. Indeed, the earliest evidence indicating that empty locations are avoided in a search task comes from a study with 23-month-olds (Mody & Carey, 2016); an age by which toddlers have a vocabulary of approximately 3-400 words and show a rich understanding of the probable outcomes of both social and physical events.

In adults, the ability to represent an array of items and empty locations in visual-spatial short term memory (VSSTM) has been investigated using the empty-cell localization task (Eriksen & Collins, 1967). In this task, an array of identical objects is briefly presented in a 4x4 grid, followed by an inter-stimulus interval and the brief presentation of another array in the same grid. The items in the two arrays taken together always fill the entire grid except for one cell, which participants have to attempt to locate. The task was originally used to study the visual integration of the two arrays presented with short inter-stimulus intervals of less than

100ms (Eriksen & Collins, 1967; Di Lollo, 1980), but by increasing the inter-stimulus interval the task targets the ability to integrate the information from the two arrays in VSSTM.

On the one hand, it has been argued that participants manage to locate the empty cell by integrating the information on the items in the two arrays into a unified representation and finding the only cell that is not filled by the collection of the items in the two arrays (Brockmole et al., 2002; Lewis, Borst & Kosslyn, 2011). Another account claims that participants form a representation of the 'negative space' in the first array and compare this representation to the negative space in the second array (Jiang, Kumar & Vickery, 2005; Hollingworth, Hyun & Zhang, 2005).

However, in this task the negative space is always the inverse of the array of items, and therefore representing the empty locations adds no further detail to the representation of items and vice versa. Furthermore, the items and empty locations in this task are always displayed in a regular grid with a visible outline, with filled items depicted as black circles almost entirely filling their cells, and empty locations depicted as grey cells. Therefore, the question of representing either 'filled' or 'empty' locations is essentially a question of representing a black pattern on a grey background or vice versa. Indeed, a study that manipulated the ratio of items and empty locations in a similar task (Chang, Chen & Yu, 2012) found that performance was worst when the two types of information were presented in a 50-50 ratio, and performance was indistinguishable when the items and empty locations were presented in a 75-25 and a 25-75 ratio. This finding suggests that as long as the law of excluded middle is met, participants only represent the information with fewer instances and use that to make inferences about the second type of information.

The current research was partially motivated by the developmental findings showing young children's failure to represent empty locations, and our central question was to explore the strategy adults might use to represent and retrieve information on empty locations. We first draw on the existing literature on the potential strategies used to represent *items* in VSSTM and then describe a novel paradigm that tests which of these strategies might be suitable for the representation of empty locations.

Firstly, an array of individual items can be encoded separately by parsing the visual scene into a set of object files which then provide information on the individuating features as well as the exact coordinate position of the object (Kahneman, Treisman & Gibbs, 1992), thus mediating between the object recognition and the spatial systems (Dent & Smyth, 2005; Irwin & Andrews, 1996; Lehnert & Zimmer, 2006). This is in line with the evidence that object information tends to activate location information and vice versa (Hommel, 1998; Olson & Marshuetz, 2005). For such high resolution information to be obtained, each object in the scene has to be attended serially, as it has been shown that visual-spatial short term memory (VSSTM) is very poor for unattended information (Rensink, O'Regan & Clark, 1997; Simons & Levin, 1998), and detail is preferentially encoded at fixations (Irwin, 1992; Henderson, Weeks & Hollingworth, 1999; Henderson & Hollingworth, 1999).

The content of the representations acquired in this way might then be independent from one another in the sense that the loss of information on an item does not automatically affect the information held on the rest of the items in the scene. However, such detailed representations are time consuming to form and accrue a high cost of maintaining them. A long-standing model suggests that up to 4 items can be held in VSSTM at any one time, which each take up an independent memory 'slot' (Luck & Vogel, 1997; Zhang & Luck, 2011) or compete for the same finite memory 'resource' (Bays & Husain, 2008; Bays, Catalao & Husain, 2009).

Secondly, an array of randomly positioned items can be represented as a global spatial configuration (henceforth *pattern*), thus expanding the capacity limitations of VSSTM. Evidence for the encoding of the relational information between items in a holistic scene comes from a study showing that the repetition of the configuration of task-irrelevant items in a visual search task aided participants' search for the target item; an effect called contextual cuing (Chun & Jiang, 1999). In addition, change detection regarding a single item is enhanced when the original context presented during the encoding phase is also shown at test (Hollingworth, 2007; Jiang, Olson & Chun, 2000; Papenmeier, Huff & Schwan, 2012; Timm & Papenmeier, 2019). Encoding individual objects as a global pattern thus facilitates the joint representation of the spatial positions of the objects in a single 'position map' (Dent, 2009).

While representing items as a pattern enlarges VSSTM capacity, and the spatial pattern can be encoded without allocating direct visual attention to each item serially (Jiang, Olson & Chun, 2000; Huang & Pashler, 2007; Yuan, Uttal & Franconeri, 2016), this representational format entails relational information between the individual items, and thus information held on one item is sensitive to the representation of the other items in the scene. Indeed, a study replicated the finding that the presentation of the entire pattern at retrieval benefits performance compared to the presentation of a single probe, but it found no memory enhancement if only 4 of the 6 original items were re-presented at test (Papenmeier, Huff & Schwan, 2012).

The two strategies of representing items individually, along with their exact positions and representing them as a global pattern incorporating a network of relational information might heavily depend on task demands, such as the number of to-be-remembered items, the length of time allowed for encoding, and whether configural information is present at retrieval. In addition, these strategies are not mutually exclusive and they might both be adopted in a single task to achieve optimal performance (Donnelly, Humphreys & Riddoch, 1991; Jiang & Wagner, 2004).

In this experiment we explored how empty locations are encoded and retained in VSSTM and whether these processes are distinguishable from the mechanisms used to represent the locations of items. Participants were presented with a set of cards at random locations on a uniform grey background, half of which the participant was required to track. Each tracked card either revealed an item or a piece of the grey background, while the content of the untracked cards remained hidden. Items were identical within a single trial, while dissimilar items were presented on each subsequent trial. The proportion of items over empty locations varied randomly with a ratio of 1:1 (medium density pattern), 1:3 (sparse pattern) and 3:1 (dense pattern). For instance, if a total of 4 cards were revealed, 3 of which were items and 1 was an empty location, this resulted in a denser pattern than the presentation of 1 item and 3 empty locations.

Following a memory delay interval, one of the tracked cards (=target) and one of the untracked cards (=distractor) turned over, resulting in a change at the target location 50% of the time. Change trials involved the deletion of an item from a previously filled location (*item/change*) or the addition of an item at an empty location (*empty/change*). During no-change trials filled target locations remained filled (*item/no change*), while empty target locations remained empty (*empty/no change*). The distractor location always complemented the target location in a way that if the target location was filled, the distractor location was always empty and vice versa. Participants were presented with 4 tracked and 4 untracked cards in the first block resulting in a smaller pattern of tracked locations, while 8 tracked and 8 untracked cards were presented in the second block resulting in a larger pattern of to-be-remembered locations. A detailed description and an illustration of the paradigm is included in the Methods section, and screen recordings of the task are attached to the Supplementary Materials.

Crucially, unlike in the empty-cell localization task, in this task it was not sufficient to represent the locations of the items in order to infer the relevant information regarding the empty locations, as an untracked card whose content was unknown to the participant was also revealed at the outcome. For instance, at the empty/change outcome an item appeared at an empty location, while at the empty/no change outcome an item appeared at an untracked location, and the participant had no means to determine whether the trial was correct or incorrect if they had ignored the empty locations during encoding.

Accurate behavioural responses in this task can rely on two strategies. Firstly, participants might serially encode the exact position coordinates of each tracked card along with their specific content. In case of the filled locations, the positions might then be wedded to a detailed representation of the surface features of the items. In case of the empty locations, however, it is not clear what type of visual information could be bound to the locations, if any. For empty locations to be represented as unique pieces of information, participants would either have to represent the *absence* of the item at a specific location, or represent that a specific location does *not* contain the item.

Secondly, participants might form a representation of the global pattern of the tracked locations and map the items onto certain parts of the pattern. This information would then aid participants in deciding whether the target card at the outcome violates the configural information held on the items or not. Crucially, this strategy does not entail any representation of the empty locations per se, but rather, it relies on the accurate preservation of the tracked locations as a global pattern and the relational information between the items in space.

These two strategies are not mutually exclusive, however, based on the task demands in this experiment, the separate encoding of each location might be the most suitable strategy. Firstly, participants were only presented with one of the tracked locations at the outcome, therefore the configural information could not aid change detection. Secondly, since a spatial pattern entails interrelated locations where each location is defined in relation to one another, encoding the information in this format is more error prone if a subject is tested on a single location. In addition, the content of half of the cards were unknown to the participant, which made it necessary to compute the exact locations of the tracked cards.

However, the lack of visual information to represent at the empty locations and the difficulty of maintaining a piece of information as a negative proposition might make it difficult, if not impossible to represent empty locations in isolation with exact position coordinates. Consequently, participants would use the most optimal strategy of retrieving information on the exact location information stored on each *item*, whereas they could only rely on the configural information involving a pattern of items when responding to changes that occur at an *empty* location.

We analysed fixation patterns and pupil dilation during encoding and memory retention, as well as change detection accuracy upon the presentation of the outcome (item/change, item/no change, empty/change, empty/no change). Our overarching prediction was that participants would opt for serial encoding as the most effective strategy when encoding items. However, we predicted that due to the difficulty of encoding the *absence* of an item at a specific location, empty locations would not be encoded and retained as unique pieces of information. In contrast, empty locations would be encoded via the parallel encoding of the global pattern of items and empty locations, and the representation of empty locations would be inseparable from the representation of the spatial pattern as a whole. A summary of our predictions is provided below and our detailed predictions with regard to each dependent measure are listed along with the presentation of the experimental methods.

Firstly, we analysed participants' fixation patterns during encoding while the content of the tracked cards was visible to establish the amount of time participants spent encoding empty and filled locations. Serial encoding relies on the assumption that each location is visited in succession and that the depth of encoding is proportionate to the amount of time spent encoding the critical location. In contrast, if individual locations are captured via parallel encoding, then successful memory retrieval is not necessarily associated with the amount of time looking at any one location. Therefore, we predicted that the time spent looking at an *item* would predict accuracy in the item/change and item/no-change conditions. However, if empty locations are encoded as a property of the global pattern, then memory for the empty location will depend on a holistic representation, and therefore it will be unrelated to the time spent encoding the empty target location.

Secondly, pupil dilation was measured during the subsequent memory delay interval while all the cards were face down and the cards remained stationary. The magnitude of pupil dilation has been linked to the number of items that the subject is required to track and maintain in memory (Kahneman & Beatty, 1966; Alnæs et al., 2014; Unsworth & Robinson, 2015). If items are encoded and retained as unique pieces of information, then increasing the number of items will be associated with increased pupil dilation. In contrast, if empty locations are exclusively represented as part of the pattern, then adding more empty locations to a pattern will not affect pupil diameter.

Lastly, change detection accuracy was measured at the end of each trial. Since this paradigm was expected to favour the strategy of coordinate position encoding and object-to-location binding, which may not be possible in case of empty locations, we predicted better performance if the target location was filled during encoding compared to when it was empty. Crucially, we predicted that empty locations would not only be harder to respond to than filled locations, but that participants would opt for radically different strategies when encoding these two types of content.

We varied the ratio of items over empty locations in a set (i.e. *pattern density*), as well as the size of the overall pattern (*pattern size*). These two manipulations generate differential predictions for the performance at filled and empty locations. On the one hand, if items are encoded via point-by-point serial encoding, then denser patterns with more items will be harder to encode than sparser patterns with fewer items. In contrast, if empty locations are encoded via the parallel encoding of the global pattern, then the memory for empty locations will be most affected by the size of the pattern, i.e. the number of tracked locations overall. In summary, pattern density was predicted to affect accuracy more when tested on an item, while pattern size was predicted to hinder performance more when tested on an empty location.

Methods

Participants

Thirty-three undergraduate and postgraduate students with normal, uncorrected vision were tested (23 females). Participants were recruited via posters displayed at [name of university]. An additional 5 participants were tested but excluded due to at chance performance in one or both of the blocks. The percentage of participants' correct responses in the two blocks is displayed in Supplementary Table 1. Participants received a small fee (£12) for taking part.

Design and Stimuli

On each trial participants saw a central fixation image accompanied by a brief auditory stimulus for 1.5s, followed by the presentation of a number of cards on a grey background (R:150, G:150, B: 150) for 2s. The locations of the cards on each trial were randomly selected from 28 possible locations on the screen. Half of these cards then lit up for 1.5s accompanied by a sound to indicate that participants were required to track the content of these cards. The highlighted cards then turned over and each of them revealed either an image or a piece of the grey background for 5s. The items at the filled locations were randomly selected from a set of 260 dissimilar images in a way that identical items appeared within a single trial, but the items differed on each subsequent trial across the experiment. The images were irregular geometric shapes edited from photographs of galaxies using Final Cut Pro. A link to the experimental stimuli is included in the Supplementary Materials.

During the encoding interval the untracked cards remained face down, and the outline of the tracked cards was not visible. The tracked cards then turned face down, and all the cards remained stationary for 5s. During this time participants were required to maintain the locations of the empty and filled cards in memory. At the outcome, one tracked card (=target) and one untracked card (=distractor) turned over, always revealing an item and an empty location. The positions of the tracked and untracked cards, as well as the locations of the target and distractor cards were randomised on each trial. Examples of the items appearing at the filled locations, the 28 possible locations on the screen and an example trial sequence are depicted in Figure 1.

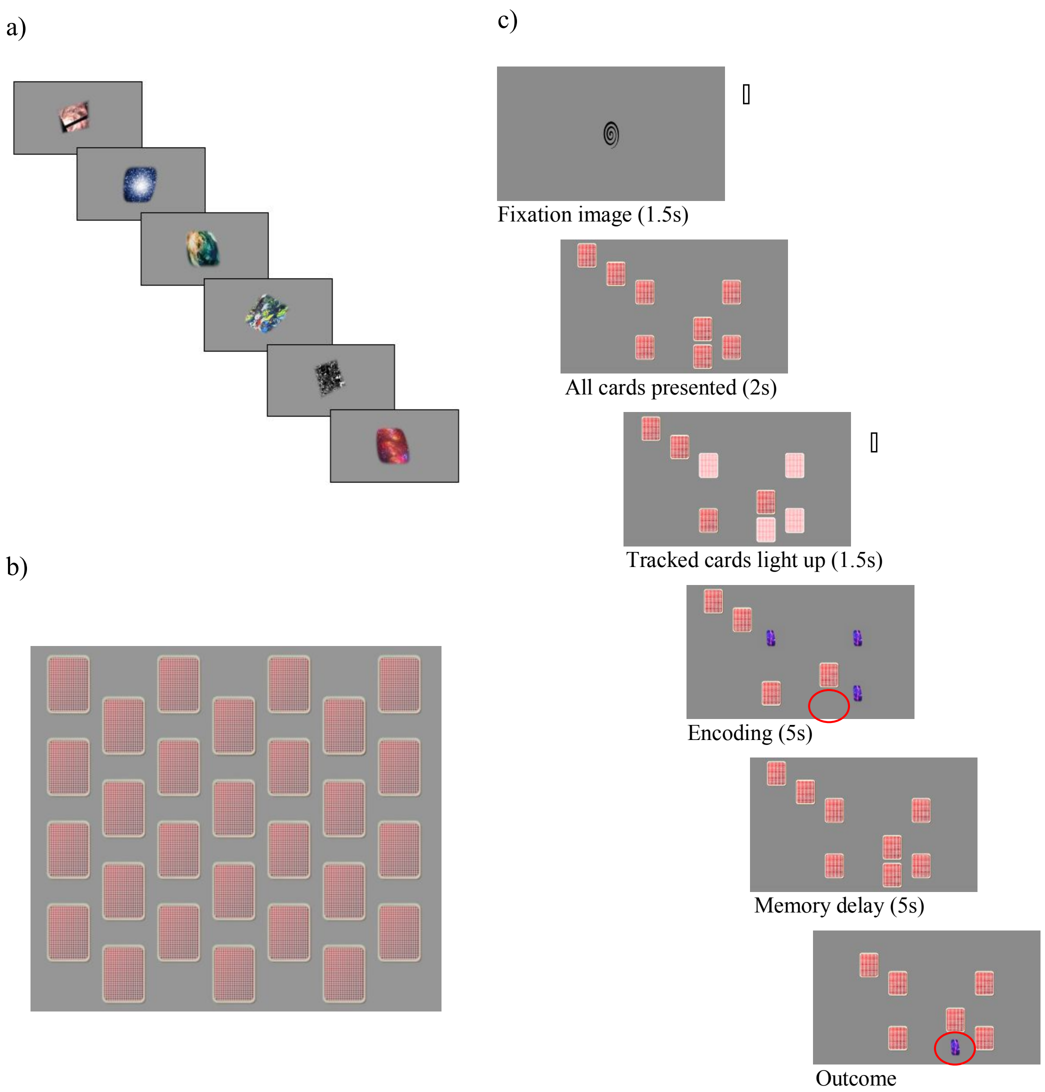

Figure 1. Experimental stimuli and example trial sequence. a) Examples of the items used throughout the trials. Identical items were presented within a single trial, while dissimilar items were used on every subsequent trial. b) The 28 possible locations on the screen. The locations of the initial set of cards were randomly selected from the possible locations c) Example trial sequence: Block1 (Small pattern) involving 4 tracked locations, in a ratio of 3 items and 1 empty location (Dense pattern), resulting in an empty/change outcome. During encoding one of the locations remained empty (circled in red), while at the outcome the location contained an item (circled in red). The red circles were not visible to the participants.

At the outcome, participants were required to indicate whether a change had occurred at the target location. A change entailed the disappearance of an item from a filled location (I-C) or the appearance of an item at an empty location (E-C), while on no-change trials filled locations remained filled (I-NC) and empty locations remained empty (E-NC). The 4 conditions (I-NC, I-C, E-NC, E-C) were presented equiprobably in a random order. On trials where the target location was filled at the outcome, the distractor location was empty and vice versa (Figure 2). The outcome remained visible until participants responded, or for a maximum duration of 3s. Trials were separated by a 1s interstimulus interval while a grey screen was presented.

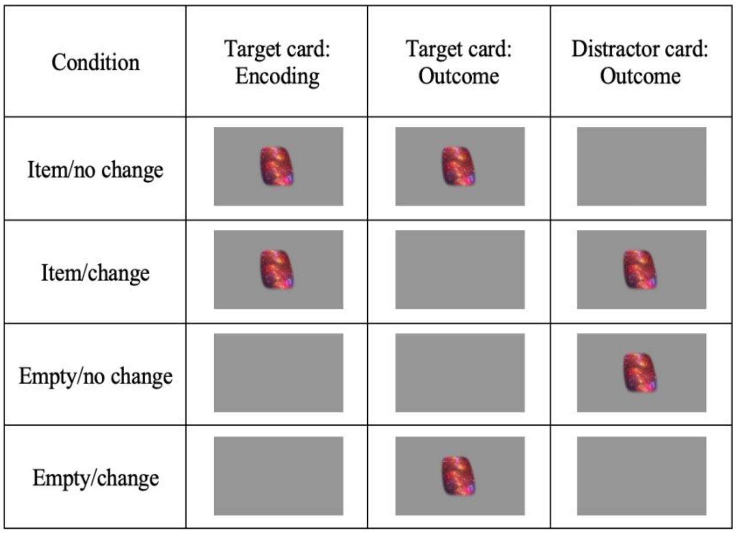

Figure 2: The content of target and distractor locations at outcome depending on the condition.
(1) Item/no change: The target location is filled both during encoding and at the outcome, while the distractor location is empty.
(2) Item/change: The target location is filled during encoding but empty at the outcome, while the distractor location reveals an item.
(3) Empty/no change: The target location is empty both during encoding and at the outcome, while the distractor location is filled.
(4) Empty/change: The target location is empty during encoding, but it reveals an item at the outcome, while the distractor location is empty.

Participants were first presented with 20 practice trials. Each practice trial consisted of two tracked and two untracked cards. One of the tracked cards was always empty, while the other one was always filled. The content of the target and distractor cards at the outcome varied as a function of condition, as described above. During the practice trials, the outcome remained visible until participants responded, and feedback was given following each trial.

Following the practice trials, participants were presented with two experimental blocks with 120 trials in each block. On each trial in the first block a total of 8 cards were presented (4 tracked and 4 untracked), while a total of 16 cards appeared on each trial in the second block (8 tracked and 8 untracked). The proportion of filled over empty locations varied with a ratio of 1:3, 1:1 and 3:1 in both blocks. In the first block participants were thus presented with a scenario of 1 item and 3 empty locations (sparse pattern), 2 items and 2 empty locations (medium density pattern), or 3 items and 1 empty location (dense pattern). Correspondingly, in the second block the tracked cards revealed either 2 items and 6 empty locations (sparse pattern), 4 items and 4 empty locations (medium density pattern), or 6 items and 2 empty locations (dense pattern). Each ratio appeared with each condition in equal proportions in a random order, resulting in 10 trials in each condition in each ratio during both blocks.

Procedure

Participants were tested in a quiet, dimmed room with constant lighting conditions (5.5 lx). Participants received explicit instructions to encode each tracked location individually, as they would be required to respond to a single card at the outcome. Participants were informed that a change could occur through an item disappearing from a previously filled location or an item appearing at a previously empty location. No-change trials were described as a filled location remaining filled or an empty location remaining empty. Participants were informed that half of the cards would not be revealed to them, and therefore the contents of these cards at the outcome were irrelevant to the task. Participants were instructed to respond as quickly and accurately as possible.

Stimuli was presented on a Tobii TX300 eye-tracker using Matlab at a sampling rate of 120Hz. Participants were seated 60cm away from the screen and used a chinrest. Responses were indicated via button press: the Right and Left arrows were used in right-handed participants and the Ctrl and Alt keys in left-handed participants (Right arrow/Alt: no change, Left arrow/Ctrl: change). After the practice trials, as well as after every 40 trials in the experimental blocks, participants took a short break (approximately 5-10 minutes). Each tracking session involving 40 trials lasted approximately 10 minutes, and the entire experiment lasted approximately 2 hours.

Data pre-processing

Gaze data: Gaze data were analysed for the 5s of the encoding phase starting with the time point when all the tracked locations were revealed, up until the point when all the cards turned face down again. Linear interpolation was applied to the gaze data corresponding to each eye with a maximum gap of 10 missing samples (83.33ms). Trials were excluded if there was no valid data for either of the eyes for at least 70% of the test interval after interpolation. Participants contributed a total of 3840 trials with valid gaze data in the first block (M = 116.36, SD = 10.19) and 3916 trials in the second block (M = 118.66, SD = 3.12). Correct responses with valid gaze data amounted to 3469 trials in the first block (M = 105.12, SD = 13.75) and 2813 trials in the second block (M = 85.24, SD = 10.65). Missing data from one eye were replaced with data from the other eye and the data were subsequently averaged across the two eyes. The area of interest (AOI) was determined as the spatial coordinates of the tracked card that was later to reveal the target location. The proportion of participants' looking time at the target location as compared to the entire scene was analysed with respect to the 5s interval while all the tracked locations were visible (Time in AOI/Valid gaze data to the entire screen during the 5s encoding interval). Target looking was compared to the participant's looking time to the entire scene because we anticipated parallel encoding in case of empty locations, which would predict scanning the global pattern of cards as opposed to the gaze dwelling in a particular AOI.

Correct trials with valid gaze data were used to assess whether participants spent more time looking at filled compared to empty targets. Correct, incorrect and missed trials were used to establish whether the proportion of target looking predicted change detection accuracy. Incorrect and missed trials were jointly interpreted as errors in that analysis.

Pupil dilation data: Pupil dilation data for correct trials were analysed for the interval while the content of the tracked cards was visible (5s) and the subsequent memory delay while all the cards were face down (5s). Linear interpolation was applied to the pupil data, as described above. Pupil baseline was established as the last 500ms of the interval while the tracked cards were highlighted, immediately before the content of the cards was revealed. Trials were excluded if there was no valid pupil data for either of the eyes for at least 70% of both the baseline and the test intervals after interpolation. Included trials were baseline corrected by subtracting the mean baseline values for each eye from the pupil data of the corresponding eye in the test intervals. Missing data from one eye were replaced with data from the other eye, and the data were subsequently averaged across the two eyes. Participants contributed a total of 3320 correct trials with valid pupil data in the first block (M = 100.6, SD = 13.32) and 2752 trials in the second block (M = 83.39, SD = 10.86).

Predictions

I. Accuracy

(i) *Condition*: Since the current task strongly favoured the representation of each tracked location, which might not be possible in case of the empty locations, impaired change detection performance was expected in the conditions where the target card was empty during encoding (E-NC, E-C) compared to when it was filled (I-NC, I-C).

(ii) *Pattern density*: If items in this task were represented individually, performance on the items (I-NC, I-C) will be impaired in the presence of denser patterns with more items. In contrast, when tested on an empty location (E-NC, E-C), participants will rely on the memory of the global pattern containing both items and empty space, therefore, increasing the number of empty locations will not impair performance.

(iii) *Pattern size*: Pattern size was expected to impair performance in case of items, owing to a greater number of items to represent in larger patterns, on average (1, 2 or 3 in small patterns and 2, 4 or 6 in large patterns). However, if the representations of items are independent from one another, then encoding for instance only 1 of 4 items might also result in a correct response when tested on the successfully encoded item. Conversely, if the memory for empty locations is embedded in the spatial relationship between the items in the pattern, then the entire pattern will have to be encoded and retrieved in order to respond to changes that occur at an empty location. Therefore, doubling the size of the pattern from 4 to 8 tracked locations is predicted to impair performance disproportionately more when responding to an empty location (empty/no change, empty/change) compared to responding to a filled location (item/no change, item/change).

II. *Gaze data*: Our primary prediction concerns the 5s encoding interval while the content of the tracked cards was visible. Longer looking to filled targets compared to empty targets on correct trials would indicate that the representation of items was formed via point-by-point serial encoding, while empty locations were preferentially encoded as part of a global pattern. Furthermore, serial encoding suggests that the amount of time spent looking at the target location during encoding will significantly predict behavioural accuracy associated with that location. In contrast, successful parallel encoding of a global pattern is not necessarily linked to dwelling in the area of the target location, therefore the association between target looking and accuracy was not predicted in case of empty locations.

III. *Pupil dilation*: Our primary prediction concerns the 5s memory delay interval while all the cards were face down and participants were required to maintain the content of the tracked cards in working memory. According to our prediction, increasing the number of items in a trial will trigger an increase in pupil dilation, indicating the amount of working memory invested into maintaining each item. In contrast, if empty space is encoded as a property of the pattern and *not* as separate entities, then the number of empty locations will not be reflected in pupil diameter.

Results

I. Accuracy

Overall accuracy was high: 90.10% in the first block and 71.61% in the second block. The percentage of correct responses in each condition in each block are displayed in Supplementary Table 2. Performance was not at ceiling or at chance in any of the conditions in either of the two blocks (see Supplementary Table 2). Accuracy was used as the primary behavioural measure, while response times showing a very similar pattern of results are reported in the Supplementary Materials. Performance in the two conditions where participants were tested on the location of an item (I-C, I-NC) was strongly correlated with performance on the empty locations (E-C, E-NC), indicating that participants attempted to maximise performance on both types of trials, $r(31) = 0.579$, $p < 0.001$. Participants contributed a total of 7920 trials (correct: N = 6404, incorrect: N = 1217, missed: N = 299).

The percentage of participants' correct responses was submitted to a 4 X 3 X 2 repeated measures ANOVA with the independent variables Condition (I-NC/I-C/E-NC/E-C), Pattern density (Sparse/Medium/Dense) and Pattern size (Block1: Small pattern/Block2: Large pattern). Bonferroni correction was applied for all pairwise comparisons. All significant findings are reported below in the order of our a priori predictions regarding Condition, Pattern density and Pattern size. Tables with the full ANOVAs conducted on the accuracy and the reaction time data are also reported in the Supplementary Materials (Supplementary Tables 3 and 4, respectively).

I.1. Condition

The ANOVA yielded a significant main effect of Condition, $F(3,96) = 36.33$, $p < 0.001$, $\eta_p^2 = 0.532$. All four conditions were significantly different from each other (Figure 3), with best change detection performance in the condition where the critical location was filled both during encoding and at the outcome (item/no change: M = 90.97, SE = 1.125), and worst performance when the critical location was empty both during encoding and at the outcome (empty/no change: M = 74.73, SE = 1.635). Furthermore, detecting the disappearance of an item from a filled location (item/change: M = 85.13%, SE = 1.66) was easier than detecting the appearance of an item at an empty location (empty/change: M = 80.37%, SE = 1.66; p = 0.030). These results confirm the prediction that encoding empty locations is significantly harder than encoding the locations of items.

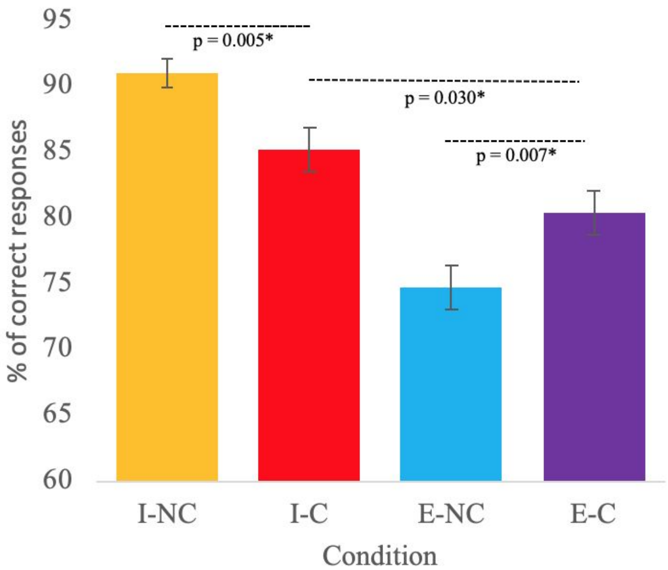

Figure 3. Percentage of correct responses in the 4 conditions collapsed across all trials. (1) Item/no change (I-NC): The target location is filled both during encoding and at the outcome; (2) Item/change (I-C): The item disappears from the target location; (3) Empty/no change: The target location is empty both during encoding and at the outcome; (4) Empty/change (E-C): An item appears at an empty location.

I.2. Pattern density

Performance in the four conditions depending on pattern density is depicted in Figure 4, and the descriptive statistics is displayed in Supplementary Table 5. In line with our predictions, a significant Condition X Pattern density interaction, $F(6,192) = 3.61$, $p = 0.002$, $\eta_p^2 = 0.101$, revealed that performance decreased linearly from sparse patterns with fewer items to dense patterns with more items if the participant was tested on an item (significant linear trends: item/no change, $F(1,32) = 30.08$, $p < 0.001$; item/change: $F(1,32) = 13.71$, $p = 0.001$). In contrast, performance on the empty locations did not show a linear trend (empty/no change: $F(1,32) = 0.47$, $p = 0.497$; empty/change: $F(1,32) = 0.48$, $p = 0.828$).

Instead, accuracy on the empty locations (E-NC, E-C) was lowest when half of the locations were filled and half of them were empty, which reached significance in the empty/no change condition (quadratic trend, $F(1,32) = 6.49$, $p = 0.016$). This result indicates the difficulty of binding the content of a particular card to a part of the pattern when both types of locations are equally frequent within the set of locations.

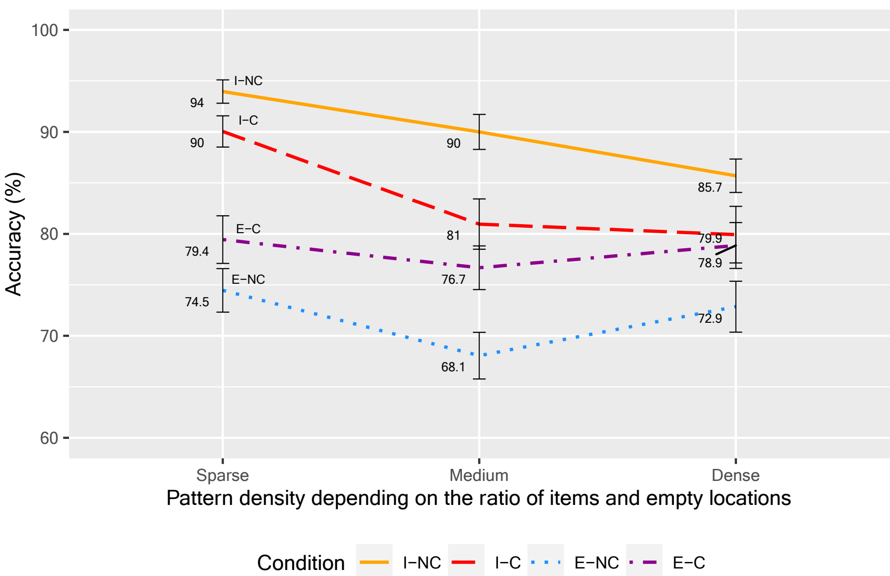

Figure 4. Changes in performance depending on the ratio of items and empty locations in the set. Sparser patterns with fewer items resulted in better performance when participants were tested on a filled location, while performance on the empty locations was not proportionate to the number of empty locations that the participant was required to represent.

To test the hypothesis that items were represented individually, planned comparisons were conducted between the scenarios that involved the same number of items in the small and large patterns (2 items + 2 empties in the small pattern compared to 2 items + 6 empties in the large pattern). Performance on the items did not differ between the small and large patterns as long as the same number of items were presented: I-NC, $t(32) = 1.21$, $p = 0.236$; I-C: $t(32) = 1.66$, $p = 0.116$. In stark contrast, presenting the same number of empty locations in a large pattern (2 items + 2 empties in the small pattern compared to 6 items + 2 empties in the large pattern) significantly impaired performance at the empty locations, indicating that empty locations were tracked as part of the global pattern: E-NC, $t(32) = 5.44$, $p < 0.001$; E-C, $t(32) = 5.56$, $p < 0.001$.

In addition, sparser patterns with fewer items were associated with better performance overall, irrespective of Condition (main effect of Pattern density, $F(2,64) = 11.86$, $p < 0.001$, $\eta_p^2 = 0.270$). This effect was likely to be driven by the item/change and item/no change conditions, as sparser patterns were easier to respond to in these two conditions, while performance at sparse and dense patterns did not differ when responding to empty locations.

These findings confirm the prediction that representing more items hinders performance when tested on an item, but empty locations do not take up independent memory resources in an additive fashion.

I.3. Pattern size

Performance in the four conditions depending on pattern size is depicted in Figure 5, and the descriptive statistics is displayed in Supplementary Table 6. As predicted, increasing the size of the pattern impaired performance in all conditions (main effect of Pattern size, $F(1,32) = 208.20$, $p < 0.001$, $\eta_p^2 = 0.867$). Crucially, this result was qualified by a significant Condition X Pattern size interaction, $F(3,96) = 10.95$, $p < 0.001$, $\eta_p^2 = 0.255$.

The largest impairment in performance (Accuracy at small patterns – Accuracy at large patterns; Figure 5b) occurred in the empty/no change condition (M = 25.96%, SE = 3.41), followed by the empty/change condition (M = 25.05%, SE = 2.50). The drop in performance from small to large patterns were comparable in these two conditions (E-NC/E-C: $t < 2$, $p >$

0.1). In contrast, performance was significantly *less* impaired in the item/change condition (M = 13.72%, SE = 2.04; E-C/I-C, $t(32) = 2.81$, $p = 0.009$). The condition that was least affected by pattern size was the item/no change condition (M = 9.11%, SE = 2.36), with no significant difference in the impairment between the item/change and the item/no change conditions (I-C/I-NC: $t < 2$, $p > 0.1$). This finding supports the hypothesis that empty locations were not encoded as individual items but as part of a global pattern containing items and empty space, and larger patterns affected performance disproportionately more in these conditions.

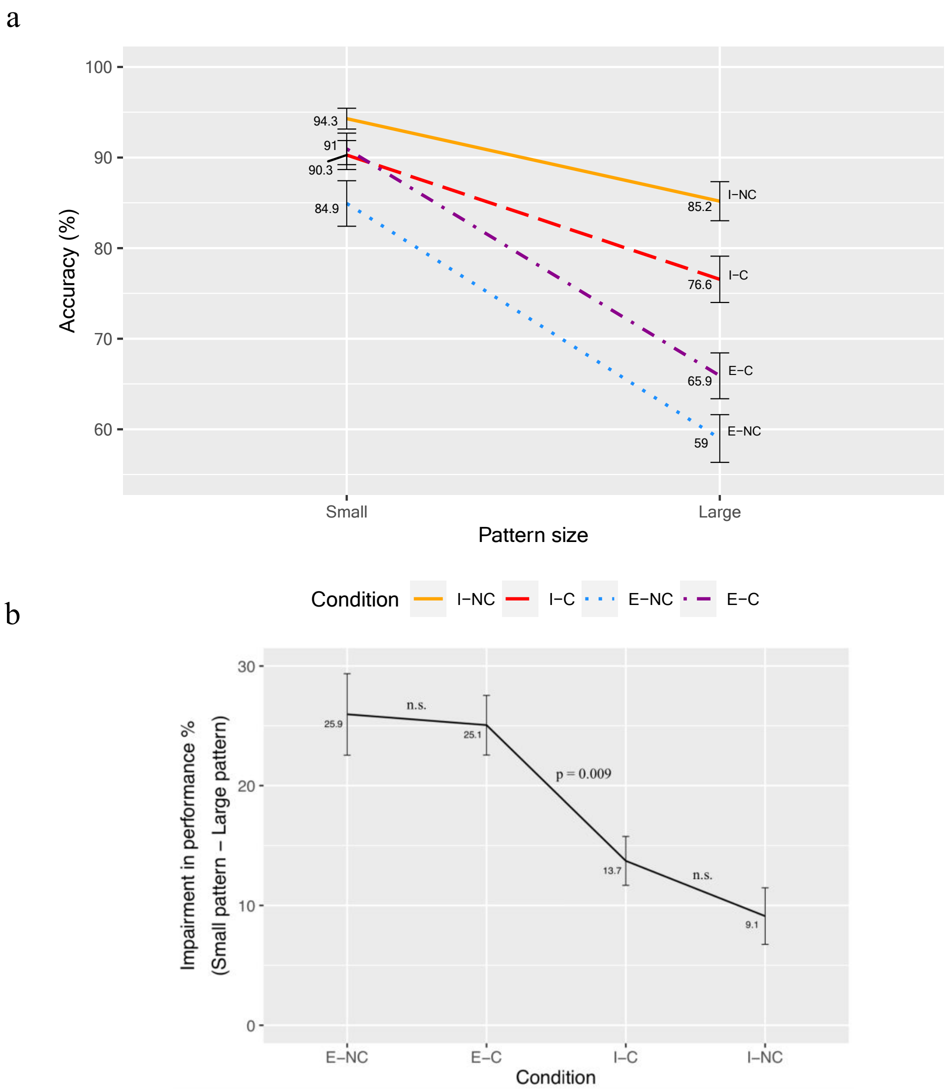

Figure 5. Changes in performance depending on pattern size. a) Performance at small patterns with 4 locations and at large patterns with 8 locations in each condition b) Impairment in performance depending on condition. Accuracy at the large pattern was subtracted from accuracy at the small pattern in each condition and compared using two-tailed t-tests.

II. Gaze data

Gaze data were analysed for the 5s encoding interval while the tracked locations were presented. As predicted, participants spent a significantly larger proportion of time encoding filled compared to empty target locations preceding correct responses in both blocks. (Block1: $t(32) = 15.39$, $p < 0.001$, filled: M = 24.52, SD = 8.11, empty: M = 3.57, SD = 1.99, Block2: $t(32) = 14.75$, $p < 0.001$, filled: M = 11.24, SD = 3.85, empty: M = 1.48, SD = 1.07). The distribution of participants' target looking time to empty and filled locations across the 5s encoding interval is displayed in the Supplementary Materials (Figures S4 and S5). The number of correct trials over all trials (Figure S4) and the proportion of correct trials over all trials (Figure S5) were calculated within 500ms time bins in case of filled and empty target locations. The graphs indicate larger gaps between the amount of correct trials over all attempted trials in case of empty locations within the same 500ms time bins. In other words, performance on empty locations was poorer than on filled locations, even when target looking times were comparable. This suggests that although empty locations were attended to for shorter periods of time on average, this fact alone cannot account for poorer memory performance in relation to this type of information.

Crucially, the time spent looking at the target during *encoding* significantly predicted subsequent accuracy if the target location was filled during encoding (Block1: $\chi^2(1) = 11.58$, $p = 0.0007$, Block2: $\chi^2(1) = 26.94$, $p < 0.0001$) but not when it was empty (Block1: $\chi^2(1) = 1.65$, $p = 0.199$, Block2: $\chi^2(1) = 0.08$, $p = 0.777$). (The data were subsetted depending on the identity of the target location (filled/empty) in each experimental block. The data in each subset was entered into a mixed effects logistic regression model. The proportion of looking in the target AOI compared to the entire scene was used as the fixed factor and participant as the random factor. *R*, package: g*lmer*).

In addition, we investigated whether these differences may have resulted from the relative salience of the filled compared to the empty locations during encoding. The data suggest that participants spent more time encoding the perceptually salient items than the vacant locations, which in turn may have influenced subsequent memory processes.

Therefore, we ran the same analyses as above, with the exception of considering only those trials where encoding times fell between 0-1000ms. Therefore, encoding times with regard to empty and filled locations were equated in this analysis. When considering the trials with short encoding times (<1s), we found the same pattern of results as above: the amount of time spent encoding the target location significantly predicted behavioural accuracy in both blocks if the target location was *filled* during encoding (Block1: $\chi^2(1) = 4.024$, $p = 0.0449$; Block2: $\chi^2(1) = 26.94$, $p < 0.00001$). In contrast, considering the same time window allocated to encoding (<1s), the time spent at the target location did not predict accuracy if the target location was *empty* during encoding (Block1: $\chi^2(1) = 0.840$, $p = 0.359$; Block2: $\chi^2(1) = 0.081$, $p = 0.777$). Consequently, the fact that empty locations were attended to for shorter periods of time cannot fully account for the subsequent differences in the depth of encoding and the likelihood of successful memory retention.

In addition, to further investigate the impact of salience, we analysed the data for the 5s *memory delay interval* while all the tracked locations were concealed. Importantly, the visual material presented during this interval was identical in all trials; participants saw a set of cards turned face down, irrespective of whether or not these cards concealed empty or filled locations. Looking at the card during the delay phase that the participant was later tested on significantly predicted accuracy if the location concealed by the card was *filled* (Block1: $\chi^2(1) = 4.622$, $p = 0.0316$; Block2: $\chi^2(1) = 17.065$, $p = 0.000036$). In contrast, the time spent looking at the card that concealed an *empty* target location did not significantly predict later accuracy. (Block1: $\chi^2(1) = 0.223$, $p = 0.636$; Block2: $\chi^2(1) = 3.496$, $p = 0.0614$).

These results jointly indicate that direct visual attention was linked to subsequent memory performance in case of filled locations. However, if participants succeeded at change detection regarding an *empty* location, their strategy was independent from allocating direct visual attention to the empty location itself.

III. Pupil dilation

Pupil dilation on correct trials was analysed for the encoding and memory retention intervals. Our primary interest was to measure cognitive processes during the memory retention interval, but pupil dilation was also analysed for the preceding encoding interval to determine whether the effects present during retention were purely a consequence of prior encoding processes.

In order to assess whether pupil dilation differed reliably as a function of the number of to-be-remembered items and empty locations, permutation analysis was conducted on the data (R, package: *permutes*, np = 1000). During the permutation analysis, the condition labels are randomly reassigned to the data numerous times at a given time point, and on each iteration a two-tailed *t*-test is conducted on this data (or an *F*-test in case of more than two conditions). The *t* values over all iterations create a distribution of test statistics values observed under the null hypothesis. The observed *t*-statistic using the actual condition labels is then compared to the distribution of *t* values derived from permutation analysis at that time point, and the difference across conditions is deemed significant if the observed *t* value falls outside the distribution of values that could have occurred by chance (i.e. a p value is significant at 0.05 if less than 50 out of 1000 *t*-tests had an absolute value larger than the one observed). Permutation testing is repeated at each data point and the resulting test statistics and significance values indicate the time points where the data could not have been obtained if the mapping between the independent and the dependent variable were random. This technique has been used in previous studies analysing pupil dilation (Kloosterman et al., 2015; Geng et al., 2015; Quirins et al., 2018; Hochmann & Papeo, 2014; Cheng, Káldy & Blazer, 2019).

Since a permutation analysis was performed at each data point throughout the intervals, multiple comparisons were corrected by adjusting the alpha level to an expected false-discovery rate of 5% using the method of Benjamini and Hochberg (1995), a technique that has been widely used to correct for multiple comparisons when analysing pupil data (Einhäuser et al., 2008; Katidioti, Borst & Taatgen, 2014; Lavín, San Martín & Rosales Jubal, 2014; Preuschoff, Hart & Einhäuser, 2011; Mill, O'Connor & Dobbins, 2016). A series of consecutive significant p values following the correction indicates a reliable difference across conditions, as well as the time course of the effect throughout the interval of interest.

Permutation analysis was performed separately on the data in the two experimental blocks (Figure 6). The black lines/top lines correspond to the time points where the uncorrected p values indicated a significant difference in pupil dilation across the 3 ratios, and the red lines/bottom lines display the p values after correcting for multiple comparisons.

The presentation of the small patterns in the first block did not elicit reliable differences in pupil dilation. In contrast, pupil diameter during the presentation of larger patterns was significantly different depending on the number of items presented. Permutation analysis indicates a significant effect during the memory delay phase starting approximately 1s after the content of the cards was concealed, which was sustained for approximately 2s while no visual changes occurred in the scene and participants were required to retain the contents of the cards in memory (6000ms-8000ms, Figure 6b).

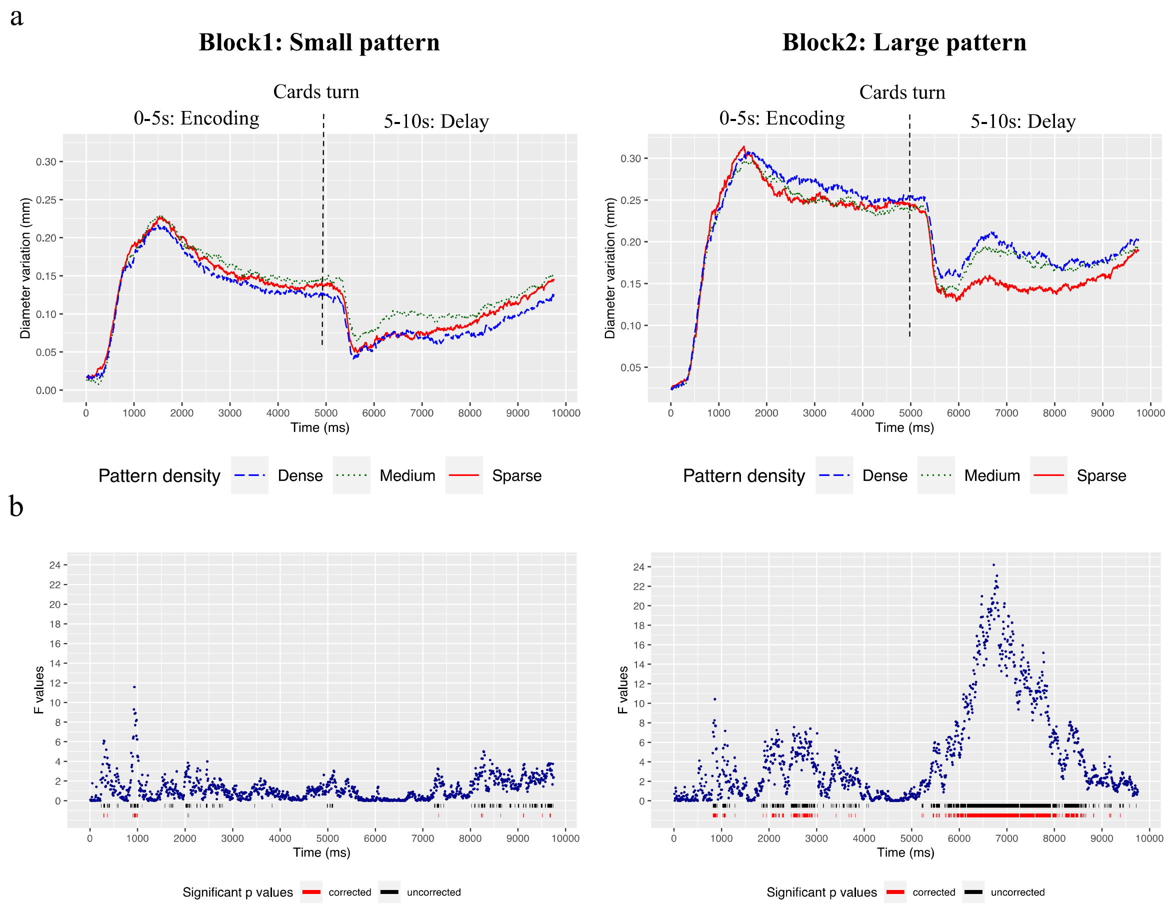

Figure 6. Changes in pupil dilation depending on the ratio of items and empty spaces at the tracked locations in the two experimental blocks. a) Changes in pupil diameter across the encoding and memory delay phases in the two blocks b) F values resulting from the permutation analyses at each time point across the encoding and memory retention intervals. The black lines/top lines depict the time points where the uncorrected p values indicate a significant difference in pupil dilation depending on pattern density ($p < 0.05$), and the red lines/bottom lines depict the significant p values after Benjamini-Hochberg correction for multiple comparisons.

In order to confirm the prediction that more items would elicit larger pupil dilation, two intervals were selected where permutation analysis indicated a significant difference: the 2000-3000ms interval during the encoding phase and the 6000-8000ms interval during the memory delay phase. Two one-way repeated measures ANOVAs were conducted on the pupil data averaged within the analysis windows in the encoding and memory delay phases with the independent variable Pattern density: Sparse (2 items + 6 empties), Medium (4 items + 4 empties) and Dense (6 items + 2 empties). All pairwise comparisons were Bonferroni corrected. Pattern density had a marginally significant effect on pupil dilation during the encoding phase, $F(2,64) = 2.84$, $p = 0.066$, $\eta_p^2 = 0.082$, with larger pupils in response to dense patterns with more items compared to sparse patterns with fewer items (p = 0.048; Dense: M = 0.271, SE = 0.022; Sparse: M = 0.249, SE = 0.022).

Crucially, during the memory delay phase while none of the tracked locations were visible, denser patterns with more items elicited significantly larger pupil dilation compared to sparser patterns, despite the fact that the same number of locations were required to retain overall. (Main effect of Pattern density, $F(2,64) = 6.90$, p = 0.002, $\eta_p^2 = 0.178$, Dense/Sparse: p = 0.005, Medium density/Sparse: p = 0.004, Medium density/Dense: p = 1.000. Dense: M = 0.179, SE = 0.025; Medium: M = 0.171, SE = 0.020, Sparse: M = 0.139, SE = 0.021.) Importantly, the differences in pupil dilation observed during the memory delay phase could not have resulted from any visual differences between the scenes during the encoding phase, as permutation analysis indicated no significant differences across ratios either at the end of the encoding phase or immediately after the tracked cards turned over (5000ms).

These results indicate that participants' pupils were significantly more dilated when retaining a larger number of items in memory, despite the fact that the overall size of the pattern was constant throughout the block. Consequently, the number of empty locations was not reflected in the magnitude of pupil dilation, indicating that items and empty locations do not compete for the same short-term memory resources.

Discussion

We tested participants' ability to encode and retrieve information regarding empty locations using a novel paradigm in which participants had to track a number of empty and filled location individually, and they could not infer the absence of an item from the presence of items elsewhere.

During encoding, participants looked significantly longer at filled compared to empty locations. Crucially, the amount of visual attention to the target location only predicted accuracy in relation to filled locations, whereas the amount of time spent looking at the empty locations was unrelated to whether or not these locations were later remembered. In other words, the impact of direct visual attention on subsequent memory was *different* depending on whether the target location was empty or filled during encoding. Follow-up analyses confirmed that this difference was present even when encoding times were equated between the items and empty locations, and persisted throughout the memory retention interval where none of the locations were visible. Therefore, the results indicate that differences in encoding and memory processes with regards to the two types of locations cannot be explained by differences in encoding opportunities alone.

Instead, the results indicate that participants adopted a strategy of encoding empty targets that did *not* involve visually attending to the relevant locations. Crucially, behavioural performance in all conditions was significantly above chance in the presence of both small and large patterns. These results are in line with previous evidence suggesting that relational information between the spatial positions of items can be encoded without allocating direct visual attention to any of the relevant items (Yuan, Uttal & Franconeri, 2016).

Secondly, in line with previous evidence, participants' pupil dilation increased with the number of items that they were required to maintain in memory (Kahneman & Beatty, 1966; Alnæs et al., 2014; Unsworth & Robinson, 2015), whereas the encoding of empty locations was not reflected in the amount of pupil dilation. These differences in pupil diameter only reached significance in the second block involving a higher memory load, and there was no further increase in dilation beyond the short-term memory capacity limit of 4 items. This is in accordance with the evidence that pupil reaches asymptote close to the critical memory threshold (Unsworth & Robinson, 2015). These findings thus provide further evidence for the hypothesis that while pupil dilation increases with the number of items held in memory, empty locations and the absence of items are not tracked as unique pieces of information and hence do not compete for the available short-term memory 'slots' (Luck & Vogel, 1997; Zhang & Luck, 2011) or resources (Bays & Husain, 2008; Bays, Catalao & Husain, 2009). From a methodological point of view, these results also indicate that pupil dilation is not simply an index of mental effort or the difficulty of encoding the stimulus, but rather it specifically signals the representational units that are held in working memory at a given time.

Thirdly, behavioural performance was significantly superior in the conditions that involved the encoding of a filled target location, which is unsurprising given the amount of prior experience representing and retrieving items compared to the rarer demand of representing empty locations.

Crucially, however, participants did not only respond less accurately to empty locations, but they had adopted a radically different strategy to encode these locations compared to the ones that contained items. Responding to items was impaired proportionately to the number of other items to encode and irrespective of the number of additional empty locations. This was confirmed by the finding that if the participant was tested on an item, then performance at 2 items and 2 empty locations was indistinguishable from the performance at 2 items and 6 empty locations.

In contrast, the same number of empty locations were associated with significantly impaired performance in larger patterns: responding to an empty location in a pattern of 2 empties and 6 items was significantly more difficult compared to the scenario of 2 empties and 2 items. These results indicate that the memory trace related to an empty location was inseparably grounded in the memory for the global pattern.

In line with previous research (Chang, Chen & Yu, 2012) we found that memory for an empty location in a large pattern of 8 locations was worst when the two types of content were equally frequent (i.e. 4 items and 4 empties). This finding suggests that participants formed a memory of the collection of all locations, and subsequently attempted to bind the less frequent content to the corresponding locations. Importantly, memory for the empty locations was embedded in the global pattern of the cards, and performance depended on retaining the location and the identity of the rest of the cards as well as the empty location.

Correspondingly, doubling the size of the overall pattern affected the memory for the empty locations disproportionately more than the memory for the items. If a set of locations was encoded as part of a global pattern, responding to an individual location might have required the retrieval of the entire pattern, as suggested by the study that found no memory enhancement if only a part of the pattern was present at retrieval (Papenmeier, Huff & Schwan, 2012). The finding that more empty locations in a pattern do not hinder performance, while increasing the size of the overall pattern results in a significant impairment supports the hypothesis that empty locations were encoded as part of the holistic representation of the scene and not as separate units of information.

Therefore, behavioural performance is closely aligned with the pupil dilation data, both lines of evidence suggesting that items *were* and empty locations *were not* tracked and retained as separate representations. In view of these results, future studies should investigate the time course and the attentional allocation associated with forming both independent partial representations as well as a holistic representation of the same scene.

In our experiment, the two strategies might have been adopted in succession within the 5s encoding interval at each trial, or the global representation may have been derived from the information held on each item, accruing no additional encoding time. However, the gaze data suggests that participants did allocate time to encode the empty locations, only that this time was not spent at the empty location itself. Average looking at a filled target location as opposed to the entire screen was proportionate to the number of tracked locations overall: 24.5% in case of 4 locations and 11.24% in case of 8 locations. In contrast, participants only spent 3.57% and 1.48% of their encoding time at an empty target location in the two blocks, respectively.

Therefore, we hypothesise that participants divided their time equally among the locations, and in case of filled locations they spent the relevant portion of the time fixating at the target, while in case of the empty locations that time was spent scanning the holistic pattern.

The results of the current experiment contribute to the interpretation of the finding that infants under approximately 2 years of age show no evidence of representing empty locations. While infants respond to the magical disappearance of an object – measured through increased looking times and brain electrical activity – they do not show a similar response when an object magically appears at an empty location (Wynn & Chiang, 1998; Kaufman, Csibra & Johnson, 2003). Indeed, even adult participants in this experiment made significantly fewer errors when detecting the deletion of an item compared to detecting the addition of an item at an empty location. These results are also compatible with the evidence that the representation of zero as a numerical value follows a different developmental trajectory compared to the representation of positive integers (Wellman & Miller, 1986; Merritt & Brannon, 2013; for a review see Nieder, 2016).

Our task strongly favoured the serial encoding of locations and their storage as separate representations. Firstly, tracked locations were presented in an array of distractors, and participants were required to retrieve accurate information regarding a single probe. Secondly, the pattern of cards provided no salient grouping cues and participants were not aided by configural information at retrieval. Lastly, participants had received explicit instructions to track each location individually. Despite that, empty locations were not encoded as unique pieces of information but merely as a property of the spatial configuration. Future studies should investigate whether it is at all possible to form such unique representations about empty locations or the absence of any item at a specific location, as in case of an absent visual object there might simply be *nothing* to represent.

The Supplementary Materials, as well as the data, the stimuli, the analysis scripts and screen recordings of the task are available on the following link: https://doi.org/10.18743/DATA.00088

The experiment was not preregistered.

References

Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, *14*(4):1, 1-20.

Baillargeon, R. (1987). Object permanence in 3½-and 4½-month-old infants. *Developmental Psychology*, *23*(5), 655-664.

Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development*, *62*(6), 1227-1246.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851-854.

Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10):7, 1-11.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289-300.

Brockmole, J. R., Wang, R. F., & Irwin, D. E. (2002). Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 315-334.

Chang, T. W., Chen, N. S., & Yu, P. T. (2012). The effects of presentation method and information density on visual search ability and working memory load. *Computers & Education*, *58*(2), 721-731.

Cheng, C., Káldy, Z., & Blaser, E. (2019). Focused attention predicts visual working memory performance in 13-month-old infants: A pupillometric study. *Developmental Cognitive Neuroscience*, *36*, 100616.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, *10*(4), 360-365.

Dent, K. (2009). Coding categorical and coordinate spatial relations in visual–spatial short-term memory. *Quarterly Journal of Experimental Psychology*, *62*(12), 2372-2387.

Dent, K., & Smyth, M. M. (2005). Verbal coding and the storage of form – position associations in visual – spatial short-term memory. *Acta Psychologica*, *120*, 113–140.

Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General*, *109*(1), 75-97.

Donnelly, N., Humphreys, G. W., & Riddoch, M. J. (1991). Parallel computation of primitive shape descriptions. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 561-570.

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences*, *105*(5), 1704-1709.

Eriksen, C. W., & Collins, J. F. (1967). Some temporal characteristics of visual pattern perception. *Journal of Experimental Psychology*, *74*, 476-484.

Geng, J. J., Blumenfeld, Z., Tyson, T. L., & Minzenberg, M. J. (2015). Pupil diameter reflects uncertainty in attentional selection during visual search. *Frontiers in Human Neuroscience*, *9*, 435, 1-14.

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly?. *Infant Behavior and Development*, *21*(2), 167-179.

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, *10*(5), 438-443.

Henderson, J. M., Weeks Jr, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 210-228.

Hochmann, J. R., & Papeo, L. (2014). The invariance problem in infancy: A pupillometry study. *Psychological Science*, *25*(11), 2038-2046.

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(1), 31-47.

Hollingworth, A., Hyun, J. S., & Zhang, W. (2005). The role of visual short-term memory in empty cell localization. *Perception & Psychophysics*, *67*(8), 1332-1343.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus – response episodes. *Visual Cognition*, *5*, 183–216.

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review*, *114*(3), 599-631.

Huang, L., Treisman, A., & Pashler, H. (2007). Characterizing the limits of human visual awareness. *Science*, *317*(5839), 823-825.

Irwin, D. E. (1992). Visual memory within and across fixations. In *Eye movements and Visual Cognition* (pp. 146-165). Springer, New York, NY.

Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and Performance XVI: Information integration in perception and communication* (pp. 125– 155). Cambridge, MA: MIT Press.

Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cuing—configuration or individual locations?. *Perception & Psychophysics*, *66*(3), 454-463.

Jiang, Y., Kumar, A., & Vickery, T. J. (2005). Integrating sequential arrays in visual short-term memory. *Experimental Psychology*, *52*(1), 39-46.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 683-702.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583-1585.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175-219.

Katidioti, I., Borst, J. P., & Taatgen, N. A. (2014). What happens when we switch tasks: Pupil dilation in multitasking. *Journal of Experimental Psychology: Applied*, *20*(4), 380, 1-17.

Kaufman, J., Csibra, G., & Johnson, M. H. (2003). Representing occluded objects in the human infant brain. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(suppl_2), S140-S143.

Kaufman, J., Csibra, G., & Johnson, M. H. (2005). Oscillatory activity in the infant brain reflects object maintenance. *Proceedings of the National Academy of Sciences*, *102*(42), 15271-15274.

Kloosterman, N. A., Meindertsma, T., van Loon, A. M., Lamme, V. A., Bonneh, Y. S., & Donner, T. H. (2015). Pupil size tracks perceptual content and surprise. *European Journal of Neuroscience*, *41*(8), 1068-1078.

Lavín, C., San Martín, R., & Rosales Jubal, E. (2014). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Frontiers in Behavioral Neuroscience*, *7*, 218,1-8.

Lehnert, G., & Zimmer, H. D. (2006). Auditory and visual spatial working memory. *Memory & Cognition*, 34 1080–1090.

Lewis, K. J., Borst, G., & Kosslyn, S. M. (2011). Integrating visual mental images and visual percepts: new evidence for depictive representations. *Psychological Research*, *75*(4), 259-271.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279-281.

Merritt, D. J., & Brannon, E. M. (2013). Nothing to it: Precursors to a zero concept in preschoolers. *Behavioural Processes*, *93*, 91-97.

Mill, R. D., O'Connor, A. R., & Dobbins, I. G. (2016). Pupil dilation during recognition memory: Isolating unexpected recognition from judgment uncertainty. *Cognition*, *154*, 81-94.

Mody, S., & Carey, S. (2016). The emergence of reasoning by the disjunctive syllogism in early childhood. *Cognition*, *154*, 40-48.

Nieder, A. (2016). Representing something out of nothing: The dawning of zero. *Trends in Cognitive Sciences*, *20*(11), 830-842.

Olson, I. R., & Marshuetz, C. (2005). Remembering "what" brings along "where" in visual working memory. *Perception & Psychophysics*, *67*, 185–194.

Papenmeier, F., Huff, M., & Schwan, S. (2012). Representation of dynamic spatial configurations in visual short-term memory. *Attention, Perception, & Psychophysics*, *74*(2), 397-415.

Preuschoff, K., Hart, B. M., & Einhauser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, *5*(115) 1-12.

Quirins, M., Marois, C., Valente, M., Seassau, M., Weiss, N., El Karoui, I., ... & Naccache, L. (2018). Conscious processing of auditory regularities induces a pupil dilation. *Scientific Reports*, *8*(1), 1-11.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368-373.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1-2), 1-46.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, *5*(4), 644-649.

Timm, J. D., & Papenmeier, F. (2019). Reorganization of spatial configurations in visual working memory. *Memory & Cognition*, *47*(8), 1469-1480.

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin & Review*, *22*(3), 757-765.

Wellman, H. M., & Miller, K. F. (1986). Thinking about nothing: Development of concepts of zero. *British Journal of Developmental Psychology*, *4*(1), 31-42.

Wynn, K., & Chiang, W. C. (1998). Limits to infants' knowledge of objects: the case of magical appearance. *Psychological Science*, *9*(6), 448-455.

Yuan, L., Uttal, D., & Franconeri, S. (2016). Are categorical spatial relations encoded by shifting visual attention between objects?. *PloS One*, *11*(10), e0163141.

Zhang, W., & Luck, S. J. (2011). The number and quality of representations in working memory. *Psychological Science*, *22*(11), 1434-1441.

Supplementary Materials

Recordings of the experimental task, the data, the stimuli and the analysis scripts are available at https://doi.org/10.18743/DATA.00088

Reaction time data

Participants contributed a total of 3568 correct trials in the first block (M = 108.12, SD = 10.03) and a total of 2836 correct trials in the second block (M = 85.93, SD = 9.94). Kolmogorov-Smirnov tests indicated that the data was normally distributed in both experimental blocks (Block1: D(33) = 0.121, p = 0.200, M = 1.49, SE = 0.037; Block2: D(33) = 0.103, p = 0.200, M = 1.55, SE = 0.032.) Response times that were more than 3 standard deviations away from the mean in the relevant experimental block were removed, which resulted in a less than 1% reduction in the data.

Response times on correct trials were submitted to a 4 X 3 X 2 repeated measures ANOVA with the independent variables Condition (I-NC/I-C/E-NC/E-C), Pattern density (Sparse/Medium density/Dense) and Pattern size (Block1: Small pattern/Block2: Large pattern). All pairwise comparisons were Bonferroni corrected. The results of the ANOVA are discussed below along with our predictions regarding Condition, Pattern density and Pattern size. The predictions are closely aligned with the predictions regarding accuracy, which is discussed in the main analysis.

I. Condition

Response times were predicted to be faster when responding to a target location that was previously filled (I-NC and I-C), compared to when it was previously empty (E-NC and E-C). Responses were predicted to be fastest when the target location was filled both during encoding and at the outcome (I-NC), and slowest when an empty target location remained empty (E-NC). Descriptive statistics of participants' response times in the four conditions are displayed in Table S2.

In line with these predictions, the ANOVA resulted in a main effect of Condition, $F(3,96) = 35.38$, $p < 0.001$, $\eta_p^2 = 0.536$ (Figure S1). Responses were fastest in the item/no change condition (M = 1070.81, SE = 37.72), which was significantly different from all other conditions ($p < 0.001$ in all three comparisons). Responses were slowest in the empty/no change condition (M = 1401.14, SE = 40.88; E-NC/I-NC: $p < 0.001$; E-NC/E-C: $p < 0.001$; E-NC/I-C: $p = 0.112$).

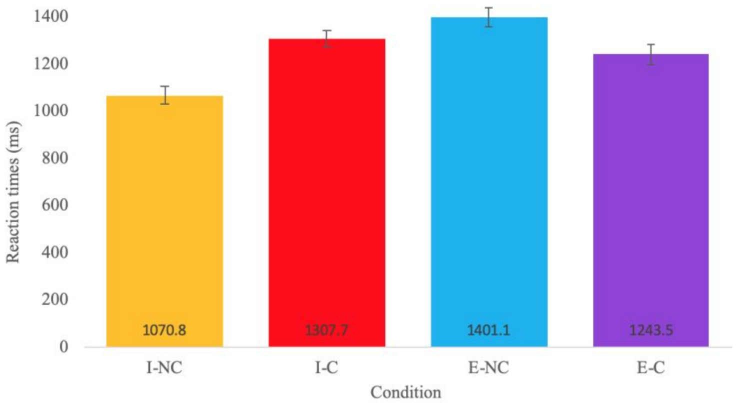

Figure S1. Reaction times on correct trials across conditions. Responses were fastest in the item/no change condition and slowest in the empty/no change condition.

In addition, we tested whether the main effect of Condition was driven by the identity of the target location during encoding (empty: E-NC, E-C; filled: I-NC, I-C) or the identity of the location at the outcome (empty: E-NC, I-C; filled: E-C, I-NC). Response times in the four conditions were submitted to a 2 x 2 ANOVA with the independent variables Encoded location (Empty/Filled) and Outcome location (Empty/Filled). This resulted in a main effect of Encoded location, with significantly slower responses if the target location was empty during encoding, $F(1,32) = 76.4$, $p < 0.001$, $\eta_p^2 = 0.705$ (Empty at encoding: M = 1669, SE = 43.5; Filled at encoding: M = 1482.49, SE = 37.49). In addition, empty outcome locations were associated with slower response times than filled outcome locations: main effect of Outcome location, *F*(32) = 9.17, p = 0.005, η_p^2 = 0.223 (Empty outcome: M = 1605.96, SE = 40.45; Filled outcome: M = 1545.68, SE = 40.41).

Importantly, a highly significant Encoded location X Outcome location interaction, *F*(1,32) = 424.55, p < 0.001, η_p^2 = 0.930, indicated that responses were fastest when the location was filled during both the encoding and at the outcome (I-NC: M = 1070, SE = 37.72) and slowest when an empty location remained empty (M = 1401.14, SE = 40.88).

These results confirm that both encoding and responding to an empty location results in slower response times compared to filled locations.

II. Pattern density

If the target location is represented as a separate unit of information, then encoding the target amongst multiple identical items will hinder performance. Namely, encoding an item will be harder in the presence of multiple other items (dense patterns), and encoding an empty location will be harder in the presence of multiple other empty locations (sparse patterns). However, according to our hypothesis, empty locations are not represented as unique pieces of information, but as part of the global pattern. Therefore, increasing the number of items in the pattern was predicted to impair performance when tested on an item (I-NC, I-C), while increasing the number of empty locations in the pattern was predicted to have a less pronounced effect when tested on an empty location (E-NC, E-C). Descriptive statistics of response times as a function of pattern density are displayed in Table S5.

A significant Condition X Pattern density interaction, *F*(6,192) = 3.85, p = 0.001, η_p^2 = 0.107, revealed that the ratio in which the two types of content were presented affected response times differently depending on the Condition (Figure S2 a). Presenting more items significantly impaired performance in the item/change and item/no change conditions (linear trends: I-NC, *F*(1,32) = 41.38, p < 0.001; I-C, *F*(1,32) = 45.05, p < 0.001). In addition, presenting more empty locations slowed down responses in the empty/change condition (linear trend: E-C, $F(1,32) = 11.06$, $p = 0.002$), while responding in the empty/no change condition was slowest when the two types of content was presented equally (quadratic trend: E-NC, $F(1,32) = 11.48$, $p = 0.002$).

The impairment resulting from encoding the target amongst multiple identical locations (i.e. an item in a dense pattern or an empty location in a sparse pattern) was compared across conditions (Figure S2 b). The largest impairment by varying the ratio of items and empty locations in the pattern occurred in the item/no change condition, followed by the item/change condition (I-NC/I-C: $t < 2$, $p > 0.1$). Performance in the empty/change condition was significantly *less* impaired (I-C/E-C: $p = 0.043$). Lastly, the ratio of items and empty locations had the smallest impact on the empty/no change condition (E-C/E-NC: $t < 2$, $p > 0.1$).

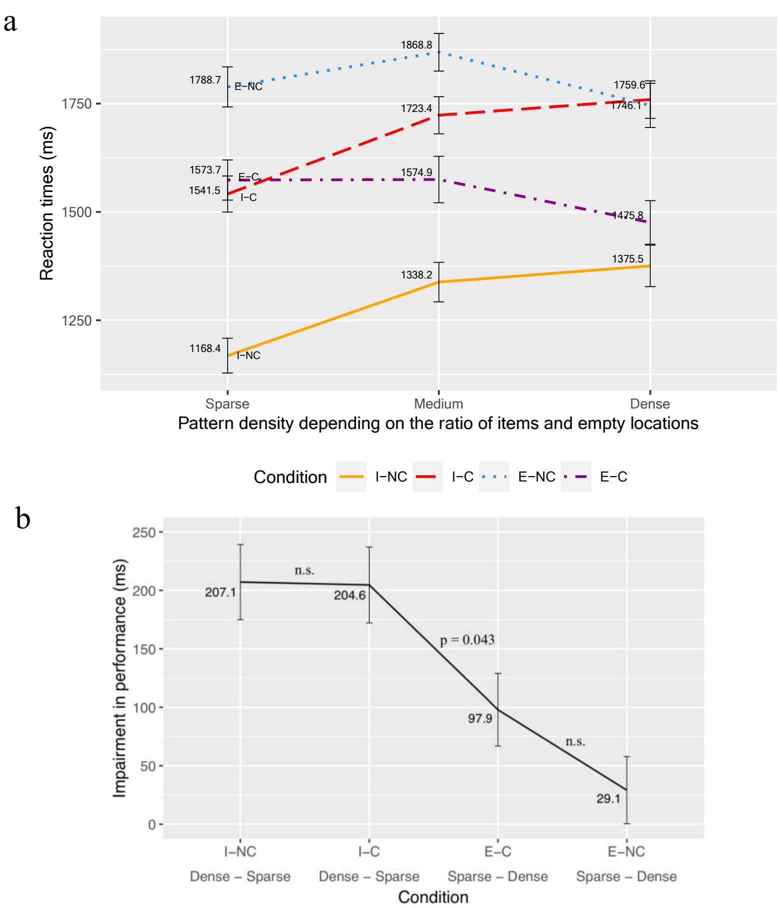

Figure S2. The impact of pattern density on reaction times in the four conditions. a) Performance at Sparse, Medium density and Sparse patterns in the four conditions. b) The impairment resulting from presenting the target amongst multiple identical locations. Difference scores were calculated by subtracting reaction times when the target was presented amongst multiple identical locations from the reaction times when the target location was more salient (Dense – Sparse when encoding items and Sparse – Dense when encoding empty locations).

Therefore, in line with our predictions, increasing the number of items in the pattern significantly impaired performance on the items (I-NC, I-C), while increasing the number of empty locations in the pattern had a less pronounced effect on reaction times (E-NC, E-C).

III. Pattern size

Similarly to our predictions regarding accuracy, the size of the overall pattern was predicted to impact performance more when the target location was empty during encoding (E-NC, E-C) compared to when it was filled (I-C, I-NC). Descriptive statistics of response times as a function of pattern size are displayed in Table S6.

The main effect of pattern size was not significant, indicating that increasing the size of the overall configuration did not have an independent effect on reaction times, $F(1,32) = 1.01$, $p = 0.325$, $\eta_p^2 = 0.030$. Crucially, however, the ANOVA resulted in a significant Condition X Pattern size interaction, $F(3,96) = 6.04$, $p = 0.001$, $\eta_p^2 = 0.159$, confirming that increasing the size of the overall pattern from 4 to 8 tracked locations had a different impact on response times depending on the condition (Figure S3).

Response times after encoding small and large patterns were not significantly different in the item/no change and item/change conditions (I-NC, $t(32) = 1.14$, $p = 0.260$; I-C, $t(32) = 0.39$, $p = 0.695$). Doubling the size of to-be-remembered locations did not affect reaction times when the participant was tested on an item, indicating that the representations of items were independent from the representations held on other locations, therefore, if the encoding of the item was successful, then responding was not slowed down by the number of other locations that had also been encoded.

In contrast, performance was significantly impaired at larger patterns if the participant was tested on an empty location (E-NC, $t(32) = 4.21$, $p < 0.001$; E-C: $t(32) = 5.54$, $p < 0.001$), suggesting that the entire pattern had to be retrieved in order to respond correctly, which in turn slowed down response times in these conditions.

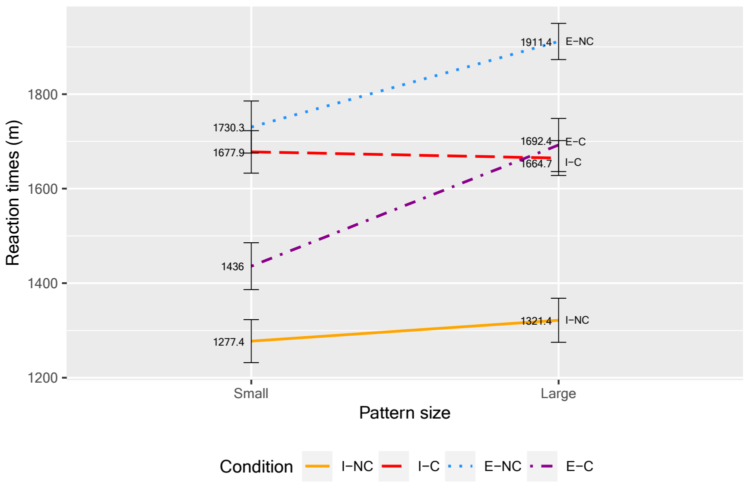

Figure S3. The effect of pattern size on the four conditions. While responses were slower after encoding large patterns when tested on an empty location (E-NC, E-C), response times were unaffected by the size of the overall pattern when tested on an item (I-NC, I-C).

Gaze data

As reported in the analysis of gaze data, participants spent significantly shorter amount of time looking at empty locations as opposed to filled locations in both blocks. Mean target looking to items was 1140.3ms in the first block (SE = 203.1ms) and 489.3ms in the second block (SE = 120.4ms). In contrast, average looking times to *empty* locations were 168.3ms in the first block (SE = 65.6ms) and 69.6ms in the second block (SE = 42.8ms). The number of correct trials over all trials (Figure S4) and the proportion of correct trials over all trials (Figure S5) were calculated within 500ms time bins in case of filled and empty target locations.

The graphs indicate larger gaps between the amount of correct trials over all attempted trials in case of empty locations within the same time windows, which suggests that poor performance on the empty locations cannot be attributed to shorter encoding times alone.

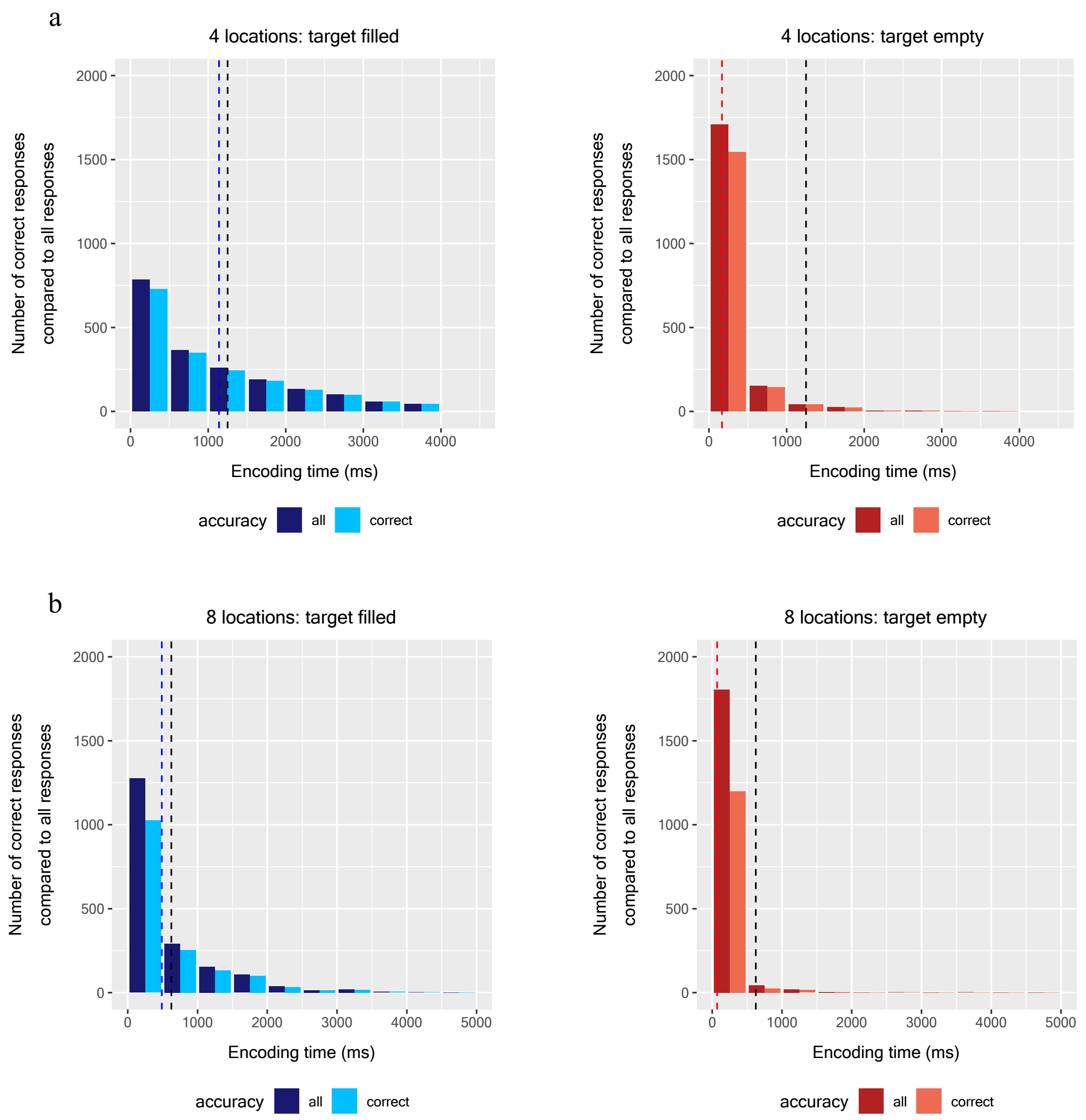

Figure S4. Total number of correct trials as compared to all trials as a function of target location, collapsed across participants. a) Block 1 (4 tracked locations) b) Block 2 (8 tracked locations). A total of 33 participants were presented with a total of 1980 filled locations and 1980 empty locations in each of the experimental blocks. (60 trials involving a filled location and 60 trials involving an empty location in each block for each participant). The dashed lines indicate the amount of time dedicated to the target location by an ideal observer (5000ms encoding time/the number of locations: 1250ms in Block1 and 625ms in Block2). The blue and red dashed lines indicate the average target looking times observed in the data with regard to filled and empty locations, respectively. (Filled target/Block1: M = 1140.3ms, SE = 203.1ms; Filled target/Block2: 489.3ms, SE = 120.4ms; Empty target/Block1: M = 168.3ms, SE = 65.6ms; Empty target/Block2: M = 69.6ms, SE = 42.8ms). Only valid trials were included in this analysis (>70% gaze data across the 5s encoding interval).

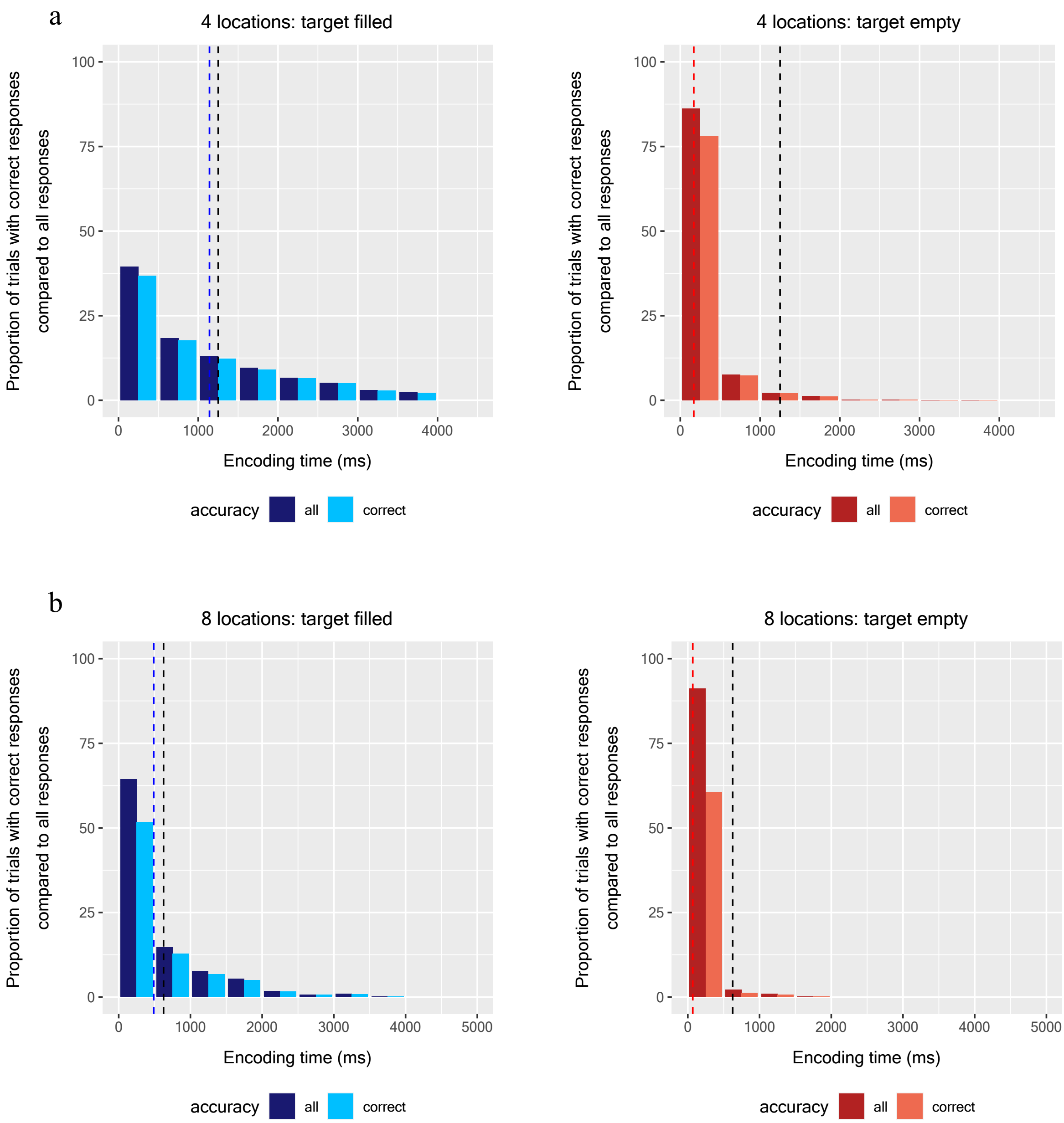

Figure S5. Proportion of correct trials as compared to all trials as a function of target location, collapsed across participants. a) Block 1 (4 tracked locations) b) Block 2 (8 tracked locations). The dashed lines indicate the amount of time dedicated to the target location by an ideal observer (5000ms encoding time/the number of locations: 1250ms in Block1 and 625ms in Block2). The blue and red dashed lines indicate the average target looking times observed in the data with regard to filled and empty locations, respectively. (Filled target/Block1: M = 1140.3ms, SE = 203.1ms; Filled target/Block2: 489.3ms, SE = 120.4ms; Empty target/Block1: M = 168.3ms, SE = 65.6ms; Empty target/Block2: M = 69.6ms, SE = 42.8ms). Only valid trials were included in this analysis (>70% gaze data across the 5s encoding interval).

Table S1. Percentage of correct trials in the training block and the two experimental blocks

ID	Training trials (% correct) Chance level: 65%	Block4 (% correct) Chance level: 57.5%	Block8 (% correct) Chance level: 57.5%
1	85	96.66	83.33
2	95	83.33	72.5
3 (excluded)	60	53.33	51.66
4	90	93.33	78.33
5	100	91.66	83.33
6 (excluded)	45	50.83	53.33
7 (excluded)	100	83.33	54.16
8	100	92.5	74.16
9	90	97.5	82.5
10	100	90	61.66
11	95	85	66.66
12	90	96.66	84.16
13	90	94.16	73.33
14	90	94.16	75.83
15	95	92.50	73.33
16	75	99.16	76.66
17	95	58.33	62.50
18	95	95	77.50
19	90	92.50	71.66
20	90	83.33	60.83
21	85	99.16	77.50
22	75	71.66	58.33
23	100	94.16	64.16
24	95	95.83	77.50
25	90	93.33	68.33
26	90	90.83	76.66
27 (excluded)	55	72.50	53.33
28	85	89.16	63.33
29	100	95.83	70.83
30	95	86.66	65
31	90	95	64.16
32	65	78.33	66.66
33	95	92.5	68.33
34	100	85	59.16
35	100	90.83	81.66
36	90	95.83	80
37	100	83.33	58.33
38 (excluded)	100	72.5	45
Mean (included participants)	**91.51**	**90.10**	**71.61**
SD (included participants)	**8.05**	**8.36**	**8.01**

Table S2. Descriptive statistics of the percentage of correct responses and reaction times as a function of Condition. Accuracy in each condition in each experimental block was compared to the test values of 100 and 50 using one-sample t-tests.

Condition Block1: Small pattern Block2: Large pattern	Accuracy (% of correct responses)	Reaction times (ms)	Test statistics comparing accuracy to 100% (ceiling)	Test statistics comparing accuracy to 50% (chance)
Block 1 I-NC	M = 94.29 SE = 1.15	M = 1077.6 SE = 45.8	$t(32) = 4.95$ $p < 0.001$	$t(32) = 38.44$ $p < 0.001$
Block 2 I-NC	M = 85.19 SE = 2.15	M = 1064.1 SE = 42.0	$t(32) = 6.87$ $p < 0.001$	$t(32) = 16.33$ $p < 0.001$
I-NC Total	**M = 90.97 SE = 1.12**	**M = 1070.8 SE = 37.7**		
Block 1 I-C	M = 90.28 SE = 1.59	M = 1356.3 SE = SE = 48.0	$t(32) = 6.07$ $p < 0.001$	$t(32) = 25.18$ $p < 0.001$
Block 2 I-C	M = 76.56 SE = 2.55	M = 1259.1 SE = 35.9	$t(32) = 9.17$ $p < 0.001$	$t(32) = 10.39$ $p < 0.001$
I-C Total	**M = 85.13 SE = 1.66**	**M = 1307.7 SE = 34.8**		
Block 1 E-NC	M = 84.93 SE = 2.51	M = 1353.2 SE = 51.3	$t(32) = 5.99$ $p < 0.001$	$t(32) = 13.89$ $p < 0.001$
Block 2 E-NC	M = 58.97 SE = 2.64	M = 1449.1 SE = 48.2	$t(32) = 15.56$ $p < 0.001$	$t(32) = 3.41$ $p = 0.002$
E-NC Total	**M = 74.73 SE = 1.63**	**M = 1401.1 SE = 40.8**		
Block 1 E-C	M = 90.96 SE = 1.74	M = 1170.6 SE = 43.2	$t(32) = 5.19$ $p < 0.001$	$t(32) = 23.53$ $p < 0.001$
Block2 E-C	M = 65.91 SE = 2.53	M = 1316.5 SE = 54.4	$t(32) = 13.47$ $p < 0.001$	$t(32) = 6.28$ $p < 0.001$
E-C Total	**M = 80.37 SE = 1.66**	**M = 1243.5 SE = 41.9**		

Table S3. Analysis of variance conducted on the accuracy data with the independent variables Condition (I-C/I-NC/E-C/E-NC) X Pattern density (Sparse/Medium/Dense) X Pattern size (Small/Large).

Effect	df	*F*	p	η_p^2
Condition	(3, 96)	36.33	<0.001*	0.532
Pattern density	(2, 64)	11.86	<0.001*	0.270
Pattern size	(1, 32)	208.20	<0.001*	0.867
Condition X Pattern density	(6, 192)	3.61	0.002*	0.101
Condition X Pattern size	(3, 96)	10.95	<0.001*	0.255
Pattern density X Pattern size	(2, 64)	1.28	0.283	0.039
Condition X Pattern density X Pattern size	(6, 192)	1.19	0.309	0.036

Table S4. Analysis of variance conducted on the reaction time data with the independent variables Condition (I-C/I-NC/E-C/E-NC) X Pattern density (Sparse/Medium/Dense) X Pattern size (Small/Large). No other effects approached significance.

Effect	df	*F*	p	η_p^2
Condition	(3, 96)	35.38	<0.001*	0.536
Pattern density	(2, 64)	0.88	0.419	0.027
Pattern size	(1, 32)	1.01	0.325	0.030
Condition X Pattern density	(6, 192)	3.85	0.001*	0.107
Condition X Pattern size	(3, 96)	6.04	0.001*	0.159
Pattern density X Pattern size	(2, 64)	1.63	0.203	0.049
Condition X Pattern density X Pattern size	(6, 192)	0.47	0.826	0.015

Table S5. Descriptive statistics of the percentage of correct responses and reaction times as a function of Pattern density. The ratio of items and empty locations was 3:1 in dense patterns, 1:1 in medium density patterns, and 1:3 in sparse patterns.

Condition	Accuracy % Sparse pattern	Reaction times (ms) Sparse pattern	Accuracy % Medium density pattern	Reaction times (ms) Medium density pattern	Accuracy % Dense pattern	Reaction times (ms) Dense pattern
I-NC	M = 93.96 SE = 1.14	M = 11.68.4 SE = 47.5	M = 90.01 SE = 1.71	M = 1338.2 SE = 45.6	M = 85.70 SE = 1.64	M = 1375.5 SE = 47.8
I-C	M = 90.04 SE = 1.52	M = 1541.5 SE = 55.2	M = 80.95 SE = 2.46	M = 1723.4 SE = 42.8	M = 79.93 SE = 2.77	M = 1746.1 SE = 51.08
E-NC	M = 74.46 SE = 2.13	M = 1788.7 SE = 54.3	M = 68.06 SE = 2.29	M = 1868.8 SE = 43.4	M = 72.86 SE = 2.49	M = 1759.6 SE = 43.3
E-C	M = 79.43 SE = 2.33	M = 1573.7 SE = 45.13	M = 76.68 SE = 2.14	M = 1574.9 SE = 53.7	M = 78.85 SE = 2.24	M = 1475.8 SE = 50.3
Total	**M = 85.56 SE = 1.13**	**M = 1518.1 SE = 38.5**	**M = 81.35 SE = 1.25**	**M = 1626.3 SE = 40.8**	**M = 81.48 SE = 1.51**	**M = 1589.2 SE = 41.9**

Table S6. Descriptive statistics of the percentage of correct responses and reaction times as a function of Pattern size. Four locations were presented in the first block and 8 locations in the second block.

Condition	Accuracy % Block 1: Small pattern	Reaction times (ms) Block 1: Small pattern	Accuracy % Block2: Large pattern	Reaction times (ms) Block2: Large pattern
I-NC	M = 94.78 SE = 1.08	M = 1277.4 SE = 45.5	M = 87.16 SE = 1.91	M = 1321.4 SE = 46.6
I-C	M = 90.99 SE = 1.46	M = 1677.9 SE = 45.04	M = 79.26 SE = 2.250	M = 1664.7 SE = 36.8
E-NC	M = 86.57 SE = 2.18	M = 1730.3 SE = 55.02	M = 62.90 SE = 2.28	M = 1911.4 SE = 38.3
E-C	M = 91.76 SE = 1.61	M = 1436.0 SE = 49.4	M = 68.97 SE = 2.36	M = 1692.4 SE = 56.2
Total	**M = 91.03** **SE = 1.30**	**M = 1530.4** **SE = 45.3**	**M = 74.58** **SE = 1.31**	**M = 1647.5** **SE = 37.49**