# Establishing individual differences in perceptual capacity

Joshua Eayrs and Nilli Lavie

Institute of Cognitive Neuroscience, University College London, UK

Address correspondence to:

Joshua Eayrs, Institute of Cognitive Neuroscience, University College London,

26 Bedford Way, London, WC1H 0AP

T: +44 (0)207 679 5372

Email: Joshua.Eayrs.13@ucl.ac.uk

Email address for Nilli Lavie: n.lavie@ucl.ac.uk

Word count : 10,010

## Abstract

Limited capacity for visual perception results in various ‘inattentional blindness’ phenomena across a wide variety of manipulations which load perception. Here we propose that these phenomena are mediated by an underlying generalised capacity for visual perception which also underlies subitizing: the ability to enumerate a limited number of items in parallel from a brief exposure. We tested this proposal by examining whether individual differences can reveal common intra-individual variance between measures of visual perception as well as of subitizing capacity. Visual perception was measured in change blindness (Rensink et al., 1997), load-Induced blindness (Macdonald & Lavie, 2008), and multiple object tracking tasks. Subitizing capacity was measured as the number of items that could be reported in parallel in an enumeration task. Perceptual capacity as indexed by subitizing was consistently a unique predictor of performance in change blindness, load-induced blindness, and motion tracking beyond any general factors that apply to both subitizing and estimation of larger set sizes. Moreover when measures of working memory were included, factor analysis indicated two orthogonal factors: perceptual and working memory. Overall, the results support the hypothesis of a generalised capacity for visual perception, and establish subitizing capacity as a predictor of individual susceptibility to inattentional blindness under load.

**Keywords:** Attention; Perceptual capacity; Subitizing; Inattentional blindness; Change blindness

## Statement of public significance

People have limited capacity for perception and will often fail to notice objects outside their focus of attention, exhibiting a form of "inattentional blindness". However, some people may have superior (while others inferior) visual detection abilities.

This study establishes a new measure that can predict a person's capacity for visual perception and object detection across multiple tasks. Results using this measure show that people who can instantly enumerate a greater number of items are found to be less prone to inattentional blindness: they can more accurately detect unattended items, changes in complex scenes, and track more moving objects compared to people with a smaller enumeration capacity.

These findings establish a new concept of generalized capacity for visual perception and awareness and provide the scientific basis for new tests that can be applied for operator screening or selection for training in many safety-critical operations (e.g. defence operators, CCTV monitoring, airport security).

# Establishing Individual Differences in Perceptual Capacity

What are the limits of our capacity to be visually aware of the world around us? This important question has stirred much interest over the years, ever since the seminal study by Neisser and colleagues (e.g. Neisser, 1979; Neisser & Becklen, 1975) demonstrated that people can fail to notice highly conspicuous events such as a woman walking with an open umbrella across their field of view when they focus attention on another event (e.g. people playing a ball game).

## Limits on Visual Perception and Awareness: “Inattentional Blindness”

Demonstrations of people’s limited capacity for visual perception come from a variety of more recent paradigms. Perhaps the most prominent of these is the ‘inattentional blindness’ paradigm. In a typical inattentional blindness task people are asked whether they noticed an unexpected and task-unrelated stimulus that is presented once at the very end of the task (e.g. Mack & Rock, 1998; Simons & Chabris, 1999; Cartwright-Finch & Lavie, 2007). Importantly, the unexpected nature of the stimulus is not critical for inattentional blindness to occur. This is shown in a ‘load-induced blindness’ (e.g. Macdonald and Lavie, 2008; Carmel, Thorne, Rees & Lavie,. 2011; Ward & Scholl, 2015) paradigm in which people are explicitly asked to detect the occasional appearance of a stimulus outside their attention focus, yet they still experience ‘blindness’ in conditions of high perceptual load in the attended task. Similarly, in the ‘change blindness’ paradigm (e.g. Rensink, O’Regan & Clark, 1997; Beck, Rees, Frith & Lavie, 2001; Lavie et al. 2014) people often fail to detect changes between stimuli across some visual disruption (for example in the form of a flicker) despite actively searching for them. The visual disruption does not interfere with the visibility of the change, but instead interferes with attention capture by the large transient signal that the change otherwise involves. In all these tasks observers are found to be strikingly unaware of stimuli and changes outside their focus of attention despite them being clearly visible. However, different task conditions can lead to different rates of inattentional blindness or change blindness and even under the same task conditions not all people suffer from inattentional blindness or change blindness to the same degree.

A major factor determining the rates of perception failures in these paradigms is the level of perceptual load in the attended task. Perceptual load has been operationally defined by referring either to the number of task units (e.g. different identity letters in a letter search task) or the level of perceptual demand (e.g. its complexity) the task requires for the same number of units (see Lavie, 1995; Lavie & Tsal 1994). Higher perceptual load so defined (for example, search tasks of larger set sizes or those requiring more complex perceptual discriminations of conjunctions of colour and shape, rather than just single colour detection) was found to result in increased rates of change blindness, inattentional blindness, and load-induced blindness for unattended stimuli(e.g. Carmel et al., 2011; Konstantinou & Lavie, 2013; Lavie, 2006; Lavie et al. 2014; Macdonald & Lavie, 2008; Remington, Cartwright-Finch & Lavie, 2014). Furthermore measures of detection sensitivity in these paradigms confirmed that 'blindness' reports reflected reduction in sensitivity of perception for unattended stimuli under high load in the task, rather than an effect of response criterion.

Importantly these findings have been replicated with a variety of perceptual load manipulations which all converge to show reduced perception accompanied by reduced visual cortex responses to unattended stimuli, under conditions of high perceptual load in the attended tasks (e.g. Rees, Frith & Lavie, 1997; Schwartz et al. 2005; Bahrami, Lavie & Rees, 2007; see Lavie, 2005; Lavie, Beck & Konstantinou 2014; Lavie & Torralbo, 2010 for reviews). Indeed in some cases high perceptual load has been shown to eliminate noticing of any task-unrelated stimuli across the entire sample (Cartwright-Finch & Lavie, 2007).

All of these findings suggest a generalised limit to the capacity for visual perception which, when fully consumed with a high load task, results in ‘blindness’ elsewhere. Crucially, perceptual load does not simply correspond to general cognitive load, but rather to specific demands on perceptual processing. Manipulations of non-perceptual ‘cognitive control’ load, for example increased working memory load, can have similar effect on task difficulty as perceptual load but lead to opposite effects on distractor processing to those of perceptual load (e.g. Lavie, 2000; De Fockert, Rees, Frith & Lavie, 2001; Lavie, Hirst & De Fockert, 2004; Carmel, Fairnie & Lavie, 2012). Thus perceptual capacity is a distinct construct from working memory and other cognitive control functions involved in the prioritizing of different task stimuli (for example in primary vs. secondary tasks, see Lavie et al. 2004; Brand D’Abrescia & Lavie 2008)

The purpose of the present study was to further characterise perceptual capacity as distinct from working memory and other more general cognitive capacities, using an individual differences approach. Thus instead of manipulating the task conditions which result in blindness across the sample we assess here the differences in the rates of blindness for a fixed level of load across different loading tasks. To test the generality of capacity limits for visual perception, we included a measure based on ‘subitizing’: the ability to accurately report the number of items from very brief display presentations (lasting a fraction of a second) in addition to more traditional visual perception tasks. We relate subitizing to visual detection performance in three different tasks: change blindness, load-induced blindness and multiple object tracking (MOT) by assessing inter-individual covariance in subitizing and detection performance in these tasks. We therefore investigate the generalised capacity for perception across diverse task demands. Establishing measures of perceptual capacity and relating these to subitizing also allows prediction of an individual’s visual detection ability, and conversely their propensity for 'inattentional blindness' by a brief and simple to administer measurement of their ability to count items from a brief display.

**Subitizing and Perceptual Capacity: Previous Research.**

Although subitizing has typically been studied within the domain of enumeration, we reasoned that it should reflect more generalized visual detection and discrimination abilities since it requires rapid detection and individuation of items simultaneously presented for a very brief duration. Indeed, in an enumeration task, responses (accuracy or reaction times) typically form a characteristic set size function consisting of two linear components, one with a very shallow or flat slope which then bifurcates at a small number of items into a serial slope as set size increases further (Kaufman, Lord, Reese & Volkmann, 1949). Encoding of sets before the bifurcation point is thus thought to be simultaneous and parallel, (as we would expect for encoding of items within perceptual capacity) serial processing then becomes necessary when the capacity for this parallel processing is exhausted (Trick & Pylyshyn, 1993; 1994). Note that although subitizing capacity is quantified by set size and reported simply as the number of items that can be subitized, this limit is not a limit on counting ability as such. Clearly people are able to count large numbers of items under less pressurizing conditions. The limit is rather in the simultaneous, parallel processing capacity for detection and individuation of items (see Piazza, Fumarola, Chinello & Melcher, 2011; Ester, Drew, Klee, Vogel & Awh, 2012).

In further support of our hypothesis, estimates of subitizing capacity across the population suggest limited capacity of circa four items, and these are similar to capacity limits estimated in previous perceptual load research. For example Lavie and colleagues have demonstrated on multiple occasions that search tasks which involve four task-relevant items or fewer do not exhaust perceptual capacity whereas larger sets do (Lavie & Cox, 1997;

Lavie & Fox, 2000; Forster & Lavie, 2008). Finally, as we briefly review below, a few recent studies of subitizing are suggestive of a link to visual perception capacity in support of our hypothesis.

Several studies have related subitizing to attention and capacity-limited visual perception using the attentional blink paradigm. In this paradigm, several stimuli are presented in rapid succession (a rapid serial visual presentation; RSVP). Stimuli presented within 700 ms of a target in this stream are less likely to be detected or recognized and this is attributed to attentional resources still being occupied by the target (Raymond, Shapiro & Arnell, 1992). When enumeration stimuli are presented in the post-target blink period, subitizing performance is significantly diminished (compared to their presentation with no preceding target; e.g. Burr, Turi & Anobile 2010; Egeth, Leonard & Palomares, 2008; Olivers & Watson, 2008; Xu & Liu, 2008).

In other attention tasks, subitizing was significantly worse when attention was directed towards a spatially (rather than temporally as in the attentional blink paradigm) separate task. For example, in a study by Railo, Koivisto, Revonuso & Hannula (2008) subitizing performance was significantly worse for unexpected dots appearing while participants were attending a line-length judgement task similar to that used by Mack & Rock (1998) to assess inattentional blindness.

Subitizing has also been shown to be significantly affected by the level of perceptual load in a concurrent task. For example, in a study by Vetter, Butterworth and Bahrami (2008), participants performed an enumeration task concurrently with a primary task of either low or high perceptual load. Their participants discriminated target shapes presented at fixation. The targets were defined by either a single feature (colour, low perceptual load) or a conjunction of features (colour and orientation, high load). Participants also attempted to enumerate the number of stimuli in a surrounding circle. High perceptual load in the central task reduced subitizing performance for the peripheral stimuli. Similarly, Chesney and Haladjian (2012) showed that while tracking multiple moving dots, the capacity for subitizing squares that were presented briefly among the dots was reduced proportionally to the number of objects being tracked. In their study, the average subitizing capacity decreased by approximately one item for every dot that was being tracked, suggesting that the same capacity limit underlies both tasks. In contrast to these findings, concurrent ‘complex span’ working memory load does not impact on subitizing performance but does have a detrimental effect on enumeration of higher quantities (Tuholski, Engle & Baylis, 2001). Thus providing further evidence that the capacity limit underlying subitizing per se (as opposed to enumeration in general) is specific to perceptual processing.

In summary, the findings of previous research are encouraging for the hypothesis that there is a general perceptual capacity limit which underlies subitizing and detection in attention-demanding task conditions. In the present research we investigate this further by assessing common variance across diverse task demands in the change blindness, load-induced blindness, MOT and subitizing paradigms.

## Study 1

The aim of Study 1 was to assess whether individual differences in detection of change between two flickering images can be predicted from perceptual capacity as measured by subitizing performance. To measure subitizing capacity, participants performed a canonical enumeration task with brief stimulus presentations. The enumeration task required participants to rapidly estimate and report the number of a briefly presented set of squares (see Figure 1). The point at which the report-accuracy/ set-size function transitioned from parallel to serial was taken as their maximal subitizing capacity. Change detection was measured using the ‘flicker task’, which is perhaps the best established measure of the phenomenon of ‘change blindness’ (Rensink, O’Regan & Clark, 1997). Participants were asked to detect the presence or absence of a change in flickering pairs of images of a real world scene (Figure 2). If a common capacity limit underlies both subitizing and change detection there should be a positive association between subitizing capacity and the ability to detect the presence of changes.

## Method

### Participants

296 participants (132 male), aged 18 to 64 (*Mean* = 31.33, SD = 13.34) volunteered to participate in Study 1. Participants did not receive any financial compensation for their time. Participants’ data were excluded from analysis if they performed at chance level accuracy in the lowest set size of the enumeration task (11 participants), or if their false alarm rate was more than two standard deviations above the group average in the change detection task (6 additional participants). The final sample analysed was therefore n = 279 (127 male), aged 18 to 64 (*Mean* = 30.40 SD = 12.07).

### Stimuli and Procedure

The data were collected in the ‘Live Science’ exhibition at the Science Museum in London over a period of several consecutive weeks. All participants provided written informed consent and had normal or corrected-to-normal vision. The study was conducted in a quiet section of the museum, this area contained three computers and so volunteers participated in groups of one to three at a time. There was always at least one experimenter present during the study; after explaining the task and obtaining informed consent the experimenter initiated the tasks. Tasks were prepared and presented in Matlab (Mathworks, Inc., Natick, MA) using the Cogent toolbox (www.vislab.ucl.ac.uk/cogent.php). Participants were seated approximately 60 cm from the screen and asked to maintain this distance but their head position was not restrained. The entire study took approximately 25 minutes to complete.

**Enumeration task.** Figure 1 illustrates a typical enumeration trial. Each trial began with the presentation of a fixation point for 1 s; this was followed by a stimulus set of black squares, presented for 100 ms each of which was randomly positioned in an area subtending 7.5cm by 7.5cm (7.15° x7.15° at a distance of 60 cm) in the centre of the screen. The squares varied in size, ranging from a minimum of 0.4cm to a maximum of 4.0cm (0.38° to 3.8° at a distance of 60 cm). The stimulus display was followed immediately by a central noise mask made up of randomly positioned black and white squares covering the same area as the stimulus display (7.5 cm by 7.5 cm). After 400 ms the mask was replaced by a central '?' to prompt a response, this remained onscreen for a further 2,400 ms or until a response was made. Participants were instructed to respond as quickly as possible indicating the number of squares displayed by pressing a key from 1-9 on the right-hand number pad of the keyboard. They could respond at any time following the initial stimulus display.

The task comprised of one practice block of 6 trials, followed by three experimental blocks of 54 trials each. After the task was explained to them, participants completed the practice block and confirmed that they understood the instructions before continuing to the experimental trials.

FIGURE 1 HERE

**Change detection task.** Each trial began with a fixation point for 1 s, which was followed by the presentation of a photograph of an outdoor scene occupying a space 22.5cm by 13.52cm (21.2° by 12.8° at a distance of 60 cm). The image was presented for 200 ms, followed by a grey rectangle of matching dimensions for 100 ms and then by a second image for a further 200 ms, which was again replaced by a grey rectangle presented for 100 ms (see Figure 2). The stimuli cycled repeatedly in this fashion for a maximum of 15 s or until participants responded. After a response was made, a green tick (3.1cm by 3.82cm) or a red cross (2.5cm by 2.5cm) appeared onscreen for 700 ms indicating that the response was correct or incorrect respectively.

The scene stimuli could either be identical (50% of trials) or could contain a slight but conspicuous change (50% of trials). Participants were instructed to respond by pressing the right shift key on the computer keyboard if a change was present or the left shift key if there was no change. They were instructed to respond as soon as they thought they knew the answer, if the 15 seconds expired with no response being made a ‘no-change’ response was recorded.

The task consisted of 44 trials in total, each of which was initiated by the participant by pressing the space bar when they were ready. After the task was explained to them, participants completed one demonstration practice and then commenced the experimental trials.

FIGURE 2 HERE

## Results

**Change detection task performance.** Table 1 presents average change detection performance for the entire sample. As can be seen in the table the false alarm rate was very low (as is typical in change blindness paradigms), a nonparametric estimate of detection sensitivity (‘A’) was therefore calculated using the formula described by Zhang & Mueller (2005). The associated measure of decision bias (b) was also calculated, these measures were used in subsequent analyses.

TABLE 1 HERE

**Enumeration task performance.** A curve fitting procedure was used to estimate individual subitizing capacity. Each participant’s accuracy (% correct) at each set size from 1-8 was fit with a bilinear function (Note 1). The function consisted of two linear components fitted to the enumeration accuracy data for each participant as follows: The function used starting values of 90% intercept and 0% slope of the first line and -15% for the slope of the second line. Each integer set size value was tested as a candidate breakpoint for the function using these starting values. The value which fit with the least error was then taken as a starting point and parameters were varied from -1 to +1 of that value using Matlab’s fminsearch function to find the best-fitting slope and intercepts. The average fit of the function to the enumeration data was good. Across the sample the average RMSE was 8.02 (SD = 4.39), the average adjusted R-squared value was 0.78 (SD = .20). Thus the bilinear function appears to predict the observed scores with little error.

The mean estimated subitizing capacity was 3.63 (SD = 0.92) which fits well with typically observed limit of three to four items in similar tasks that involve very brief display durations (e.g. Burr, Turi & Anobile, 2010). As a measure of performance not dependent upon subitizing capacity (i.e. 'estimation' performance), the average accuracy across all set sizes above the bifurcation point was calculated for each participant. The sample mean estimation accuracy was 46.02% (SD = 10.91).

**Detection rate.** Table 2 presents the correlation matrix for all the variables measured in Study 1. Subitizing capacity was significantly correlated with change detection rate ($r(278) = .31$, $p < .001$): Individuals who could subitize more items were more likely to accurately detect changes. There was also a correlation between change detection rate and estimation accuracy ($r(278) = .15$, $p = .013$), which was significantly weaker than that between subitizing and detection rate (difference z-score = -2.24, $p = .03$; calculated using the method described by Hittner, May & Silver, 2003 and implemented using the cocor package in R; Diedenhofen & Musch, 2015).

TABLE 2 HERE

A hierarchical regression was used to examine the unique contribution of subitizing capacity to change detection when including estimation accuracy to control for general factors. The regression included two steps, the first of which included only estimation accuracy as a predictor of detection rate, subitizing capacity was then added to the model in a second step. Both steps were significant (Step 1: Adjusted $R^2 = .02$, $p = .013$; Step 2: Adjusted $R^2 = .14$, $p < .001$) and both subitizing capacity and estimation accuracy accounted for a significant portion of unique variance in the final model (as shown in Table 3). Any common variance associated with task-general factors (such as motivation), would be expected to affect both measures similarly. By establishing a significant unique contribution of each predictor we therefore establish that general factors do not account for the relationship with change detection.

TABLE 3 HERE

**Detection sensitivity.** In order to establish that the association between tasks was not due to differences in response bias, the analyses were replicated using detection sensitivity (A; Zhang & Mueller, 2005). Change detection sensitivity was positively correlated with subitizing capacity ($r(278) = .37$, $p < .001$). In contrast, detection decision criterion (b) was not correlated with subitizing capacity ($r(278) = -.08$, n.s.). These findings support the hypothesis that subitizing can predict visual perceptual capacity rather than response criterion. Again, there was also a positive correlation between change detection sensitivity and average estimation accuracy ($r(278) = .15$, $p = .013$) which was significantly weaker than the correlation between subitizing and detection sensitivity (difference z-score = -3.12, $p = .002$).

FIGURE 3 AND FIGURE 4 HERE

Hierarchical regression was run, using subitizing capacity this time to predict change detection sensitivity, while controlling for estimation accuracy as before. Both steps of the regression were significant (Step 1: Adjusted $R^2 = .02$, $p = .013$; Step 2: Adjusted $R^2 = .19$, $p < .001$) and both subitizing capacity and estimation accuracy accounted for a significant portion of unique variance in the final model (as shown in Table 4), replicating the detection rate findings.

TABLE 4 HERE

The results of Study 1 support the hypothesis that perceptual capacity as measured by subitizing can predict the rates of change detection or blindness. These results held when predicting either detection rates or detection sensitivity. Thus, increased capacity to subitize is associated with better ability to accurately detect changes and not simply an effect of response criterion. This relationship is unlikely to be explained simply by general factors such as motivation due to the fact that subitizing accounts for a significant portion of unique variance when such factors are controlled by including estimation accuracy in multiple regression analyses. Larger number estimation is thought to reflect a different cognitive process to subitizing (e.g. Burr, Turi & Anobile, 2011; Vetter, Butterworth & Bahrami, 2011; Cutini et al. 2014) yet should have also been affected by general factors such as motivation. The finding of unique variance explained by subitizing capacity when estimation accuracy is controlled suggests that the perceptual capacity underlying subitizing is a specific predictor of change detection. The significant association between estimation accuracy and detection sensitivity was less expected. It is possible that the relationship may be related to some other cognitive process to which Study 2 may provide insight.

## Study 2

Study 1 established that individuals with greater subitizing capacity are better able to detect changes in the flicker change detection paradigm, indicating a greater capacity for visual perception. In Study 2 we further investigated whether subitizing capacity is predictive of visual detection in a modified ‘load-induced blindness’ task (Macdonald & Lavie, 2008). Participants performed a central task in which they made a line-length judgement of a centrally presented cross, the difficulty of which was established to be of an intermediate level of load in prior research (Remington, Cartwright-Finch & Lavie, 2014). While performing the central task, participants attempted to detect the presence of a contrast increment in peripherally presented grating stimuli (see Figure 5). Our hypothesis of a generalized capacity for visual perception leads to the prediction that there should be a positive association between individual’s subitizing capacity and the ability to detect the presence of a contrast increment in the visual periphery.

## Method

### Participants

The participants were visitors to the Science Museum that approached the researchers on the museum ‘Live Science’ exhibition floor volunteering to take part in the study. 165 participants (80 male) aged 18 to 62 (*Mean* = 26.56, SD = 9.75) completed Study 2. Participants’ data were excluded from analysis if they performed at chance level accuracy in the lowest set size of the enumeration task (10 participants), if their detection sensitivity was more than two standard deviations below the group average in the control block of the load-induced blindness task (13 participants) or if their accuracy in the central cross arm judgement was near chance in the main block of the load-induced blindness task (20 participants). The final sample analysed was therefore n = 122 (59 male) aged 18 to 62 (*Mean* = 25.22, SD = 8.89).

As in Study 1 the sample size was dependent upon the number of museum visitors who were interested in taking part. Our sample of n = 122 provides a power greater than .99 assuming a similar effect size to that in Study 1 (i.e. Cohen's $f^2$ = .21 for an $R^2$ change of .17 when predicting change detection sensitivity).

**Stimuli and procedure**

The study was run in the Science Museum 'Live Science' exhibition over a period of several consecutive weeks which followed the run of Study 1. The study was run under the same conditions as Study 1, in groups of one to three with at least one experimenter present at all times. The experimenters explained the tasks to the participants and then initiated the study after obtaining written informed consent. Once again volunteers received no payment for their participation. Participants completed the same visual enumeration task as described in Study 1. Instead of the change detection flicker task, this time participants also completed a load-induced blindness task. The entire study took approximately 25 minutes to complete.

**Load-induced blindness task.** Each trial started with a central fixation dot presented for 1 s. This was followed by a central cross shape and four peripheral back and white square gratings for 120 ms. A blank screen was then presented for 1880 ms, followed by a central '?' for 100 ms and then another blank screen for a further 1900 ms. The cross shape was formed of one vertical and one horizontal line, one of which was always longer (4.5 cm; 4.1° at a distance of 60 cm) than the other (3.5cm; 3.3° at a distance of 60 cm). On a randomly selected 50% of trials the vertical arm was longer; on other trials the horizontal arm was longer. Each square grating was 3.6 cm x 3.6 cm (3.4° x 3.4° at a distance of 60 cm) and was presented in one of the four display corners 6.4 cm (6.1°) from the nearest grating edge to the centre of the screen (extending to a maximum of 10.9° into the periphery). The contrast of the (non-target) gratings was 10%; on 25% of trials the contrast of one (target) grating was incremented by an additional 28%.

Participants were instructed to respond immediately after the stimulus presentation by pressing either the up arrow key or the left arrow key to indicate which cross arm (vertical or horizontal) was longer. A 1,900 ms blank interval followed the task presentation and this interval elapsed irrespective of the participant's response latency. A central question mark symbol '?' was then presented and participants were instructed to indicate whether the contrast increment was present in any of the four gratings by pressing the spacebar upon seeing this display

The task included one practice block of 10 trials, followed by two experimental blocks of 32 trials and finally one control block of 32 trials. The control block was identical to the experimental blocks except that participants were not required to attend to arm length discrimination. In the control block the '?' prompt appeared immediately after the stimulus display and participants were instructed to respond by indicating whether a contrast increment was present in one of the gratings. The control block was used to ensure visibility did not play a role in detection performance. Participants with chance level of performance on this block were excluded from the main analysis as described in the participants section. Participants were instructed to maintain fixation throughout the task as no one grating was more or less likely to be brighter than the others.

FIGURE 5 HERE

## Results

**Load-induced blindness task performance.** Table 5 presents average performance in the contrast detection task for the entire sample. The same measure of detection sensitivity (A) and response criterion (b) as were used in Study 1 were calculated for this data and used in subsequent analyses.

TABLE 5 HERE

**Enumeration task performance.** As in Study 1, individual subitizing capacity was estimated by fitting a bilinear function to the individual accuracy data at each set size (excluding set size nine). Again, the bilinear function fit the data well (Average RMSE = 8.10, SD = 4.40; average R-squared = .77, SD = .20). Average subitizing capacity was 3.32 (SD = 0.83) which again fits well with the typically observed limit of three to four items. Average estimation accuracy was 51.03% (SD = 11.01).

**Detection rate.** Table 6 presents a matrix of the correlations between performance measures of each of the tasks in Study 2**.** Subitizing capacity was positively correlated with detection rate in the load-induced blindness task ($r(121) = .29$, $p < .001$), as was estimation accuracy ($r(121) = .22$, $p = .014$ and these correlations were not significantly different ($z = -0.65$, $p = .51$).

TABLE 6 HERE

As in Study 1, hierarchical regression was used to measure the unique contribution of each variable. The regression consisted of two steps, the first step included only estimation accuracy and the second step included estimation accuracy followed by subitizing capacity. Both steps of the regression were significant (Step 1: Adjusted $R^2$= .04, p = .014; Step 2: Adjusted $R^2$ = .09, p = .008), as can be seen from Table 7 in the final model subitizing capacity accounted for a significant portion of unique variance, while estimation accuracy did not. As in Study 1, common variance associated with general factors such as motivation would be expected to affect both subitizing and estimation performance similarly. The finding that subitizing alone is a significant unique predictor of stimulus detection provides evidence against alternative accounts in terms of any general task performance factors.

TABLE 7 HERE

**Detection sensitivity.** Detection sensitivity was positively correlated with subitizing capacity (r(121) = .38, p < .001). Detection decision criterion (b) was not significantly correlated (r(121) = .08, n.s.). Estimation accuracy was also correlated with detection sensitivity (r(121) = .25, p = .006) and although the correlation was numerically smaller than that with subitizing capacity the two were not significantly different (difference z-score = -1.25, p = .21).

FIGURE 6 AND FIGURE 7 HERE

Once again a hierarchical regression was used to measure the unique contribution of each variable as in the analysis of detection rate. Both steps of the regression were significant (Step 1: Adjusted $R^2 = .05$, $p = .006$; Step 2: Adjusted $R^2 = .15$, $p < .001$), as can be seen from Table 8 in the final model subitizing capacity accounted for a significant portion of unique variance while estimation accuracy did not. As in Study 1, the finding that subitizing alone is a significant unique predictor of stimulus detection provides evidence against alternative accounts in terms of any general task performance factors such as motivation since these would be expected to affect both subitizing and estimation performance similarly

TABLE 8 HERE

The results of Study 2 replicate the relationship between subitizing capacity and visual detection abilities established in Study 1. The positive association between these measures indicates a common underlying resource, one which cannot be attributed to general factors such as motivation, as demonstrated by the unique variance accounted for by subitizing. Although estimation accuracy in this study was positively correlated with stimulus detection this correlation was not significant in a multiple regression including subitizing capacity, suggesting estimation accuracy alone was not a predictor of detection.

The dissociation between the correlation of estimation accuracy with detection sensitivity in change detection, but not load-induced blindness tasks is potentially interesting. One may speculate that the association between large set size estimation and detection of changes in a meaningful visual scene is related to the ability to extract the ‘gist’ (or summary statistic) of a visual display (e.g. Alvarez & Oliva, 2008). Outside of focused attention, the numerical gist of large quantities or the gist of a change versus ‘no-change’ image may benefit from the same cognitive process. Whereas the local contrast increment detection required in the load-induced blindness task may depend on a more precise level of representation that can only be obtained within perceptual capacity, thus accounting for the selective correlation with subitizing capacity but not estimation of larger set sizes (see Ward, Bear & Scholl, 2016 for a relevant discussion). As this is mere speculation at present and the focus of the present study was on subitizing rather than estimation we do not dwell further on this.

## Study 3

Studies 1-2 established subitizing as a predictor of visual detection in the change detection and load-induced blindness paradigms. In both studies subitizing was a significant predictor when controlling for larger number estimation. This provides some evidence that an individual’s subitizing capacity represents a distinct predictor of visual detection that is unlikely to reflect general cognitive factors or strategies since these would apply to the estimation performance for larger numbers. In Study 3 we sought to further assess our hypothesis that subitizing capacity is reflective of a generalized perceptual capacity as distinct from general cognitive ability by specifically examining the relation to working memory capacity. Working memory is a well-established predictor of individual differences in a range of attention tasks, including for example the Stroop task, spatial cuing and task switching (e.g. Kane & Engle, 2001; Kane, Bleckley, Conway & Engle, 2001; Redick & Engle, 2006). However as discussed earlier (in the General Introduction) in load theory perceptual capacity and working memory capacity are two dissociable functions, and there are numerous demonstrations of opposite effects of working memory load and perceptual load on attention in support of this claim (see Lavie et al. 2004 for review). Thus, if subitizing reflects perceptual capacity it should remain a significant unique predictor of visual perception when controlling for individual differences in working memory capacity. In order to specifically address the cognitive control aspect of the working memory system (Note 3) we chose three complex span tasks that involve not only memory retention but also load cognitive control in the form of multiple task demands. These have reliably been shown to predict performance of tasks requiring cognitive control of attention (e.g. Redick & Engle, 2006) and are thus predicted by load theory to be dissociable from perceptual capacity. Furthermore, recent research has demonstrated that using multiple versions of these tasks, which involve different stimulus categories (such as numerical, spatial and verbal memoranda), can produce measures more sensitive to domain-general working memory functions than when using only a single task (e.g. Foster et al. 2015; Oswald et al. 2015).

Furthermore, in Studies 1-2 the enumeration task and both visual detection tasks required detection of briefly presented (less than a quarter of a second) or transient stimuli. It is plausible that the common variance is restricted to the capacity for perception from rapid, transient presentations rather than perceptual capacity in a wider sense which extends to longer and more continuous presentations. To test this in Study 3 we included a continuous measure of perceptual capacity (MOT), which does not involve rapid or transient displays. In Study 3 we thus assessed individual differences in working memory capacity, subitizing, transient change detection in the flicker task (as used in Study 1) and continuous perception of moving objects with the MOT task.

## Method

### Participants

72 (43 female) participants aged 18 to 52 (mean = 24.42, SD = 6.99) were recruited from the UCL psychology research volunteer database and each received £7.50 for their time. All participants provided written informed consent prior to taking part. As in Study 1,

participants were excluded if their accuracy was near chance for the lowest set size of the enumeration task leading to the exclusion of four participants (three male) and a final sample size of n = 68. Assuming a similar effect size to Studies 1 and 2 (i.e. Cohen's $f^2$ = .21, see Study 2) the sample of n= 68 provided a power of .96 to detect a change in $R^2$ in a regression controlling for estimation accuracy and three measures of working memory capacity (see procedure).

## Procedure

Study 3 was run at UCL, in a quiet testing room and volunteers participated one at a time. Participants completed a total of five tasks: three complex span working memory tasks, an enumeration task, an object tracking task (MOT) and a change detection flicker task. There was always a researcher present in the room with the participants during the study. The study took approximately one hour.

**Enumeration task.** Participants completed the same enumeration task as in Studies 1 and 2 except for the following changes: The length of each block was increased to 81 trials and a fourth experimental block was added to the task (producing a total of 324 trials).

**Change detection task.** Participants completed the same change detection task as used in Study 1, however this time the task included 8 additional trials (total = 52 trials). The task structure was the same except that the flickering presentation time was reduced to 8 seconds.

**MOT task**. Participants completed an MOT task in which they were required to track four target dots as they moved around the centre of the screen among four non-target dots. The dots subtended 0.5 cm by 0.5 cm (0.5° at a distance of 60 cm) and moved randomly within an area subtending 6cm by 6cm (5.72° at a distance of 60 cm) at the centre of the screen. On each trial, eight black dots were presented against a grey background, after 500 ms four of the dots became blue for 1.5 s, after which they returned to being black. After another 500 ms the dots began to move at a rate of 2.15 degrees per second, the dots bounced off one-another and off the edges of the movement area. After 8 seconds the dots ceased movement and a single probe dot became blue once more. Participants responded to the probe by pressing the '1' key on the keyboard number pad if the probe was a target and the '2' key if not. The probe then turned either green or red to indicate a correct or incorrect response respectively. A fixation cross was then presented for 1 s before the next trial started.

FIGURE 8 HERE

**Complex span tasks**. Participants completed three complex span working memory tasks: the 'Operation Span' (OSPAN) task, the 'Reading Span' (RSPAN) task and the 'Symmetry Span' (SSPAN) task (the same as those described in Oswald et al., 2015). These shortened versions of the tasks provide the opportunity for a more representative measure of working memory – nonspecific to a particular cognitive modality (numerical, lexical or spatial). Using a variety of shortened working memory span tasks in lieu of a single full length task has been shown to provide a better measure of underlying capacity (See Oswald et al. 2015 and Foster et al. 2015 for an in-depth discussion).

E-Prime 2.0 was used to run these tasks. The task procedure was similar for all three tasks. In the OSPAN task, participants were presented with a series of sums (e.g. (8/2)+9 = 13). They responded to each sum by clicking 'yes' or 'no' icons on the screen to indicate if the given answer was correct. Following each sum, participants were presented with a letter which they memorized. After a variable number of sum and letter presentations (4-6) the participant was presented with a memory response screen in which they selected the memorized letters in the order in which they were presented. If they were uncertain of a given memoranda they responded with ‘blank’. The number of letters recalled in the correct sequential position over the course of the task provided the participant’s ‘span’ score.

Before starting the task the participants completed a series of practice trials. First only performing the ‘sum’ component with no memoranda, then only the ‘memory’ component with no sums and finally both together as in the experimental trials. The average reaction time from the final practice section plus two standard deviations was used as the time limit in experimental trials. Throughout the task the participant’s accuracy (% correct) in the operation portion of the task was displayed, and they were instructed to maintain a minimum of 80% accuracy.

The RSPAN task followed the same task structure, however instead of a sum on each trial the participant read a sentence and responded to indicate whether or not the sentence made sense (e.g. ‘The prosecutor’s dish was lost because it was not based on fact.’). The SSPAN task followed the same structure also, but instead of a sum or sentence, the participant responded to a black and white block image, indicating whether or not the left and right sides were mirror symmetrical. In the SSPAN task, instead of memorizing letters, the participant memorized the position of a black square in a white grid.

FIGURE 9 HERE

## Results

**Enumeration task performance.** A bilinear function was fitted the individual accuracy data at each set size (excluding set size nine) as before, in order to estimate individual subitizing capacity. Again, the bilinear function fit the data well. The same curve fitting procedure was used as in Studies 1- 2 to estimate individual subitizing capacity. The average RMSE for the fit was 8.26 (SD = 4.75), average adjusted R-squared was .76 (SD = .22), indicating a good fit of the model to the data. As previously, subitizing range was estimated for each participant based on the point at which the two linear components of the bilinear function intersected. Average subitizing capacity was 3.17 (SD = 0.84) which is within range of our previous findings using this task. Accuracy at set sizes beyond the subitizing range was averaged as a measure of 'estimation' ability, average estimation accuracy was 49.10 (SD = 14.35), which was also within the range found earlier.

**Change detection task performance.** Table 9 presents average performance for the change detection task in Study 3; once again false alarm rates were very low so the same non-parametric measure of detection sensitivity (A) was calculated with the corresponding measure of bias.

TABLE 9 HERE

**MOT task performance.** Average accuracy in the MOT task was well above chance (mean accuracy = 75.19%, SD = 12.76). In order to obtain an estimate of tracking capacity comparable to the subitizing capacity estimate we calculated the 'Effective Number of Objects Tracked' (ENOT) using the formula described by Scholl, Pylyshyn and Feldman

(2001). ENOT scores are calculated as m = n(2p-1), where m is the estimated tracking capacity (ENOT), n is the number of tracking targets and p is the proportion of correct responses. The average capacity based on this formula was 1.90 (SD = 0.84).

**Complex span task performance.** Table 10 presents average performance data for each of the complex span tasks. The 'total' score on each task was used as this has previously been established as the better measure of individual capacity (Redick et al. 2012). The total score is calculated as the total number of memoranda (letters or square positions) reported in the correct position in sequence (ignoring incorrect or unknown items). SSPAN score is necessarily lower as there were fewer overall trials in this task: the maximum possible scores are 30 for the OSPAN and RSPAN tasks and 24 for the SSPAN task. As can be seen from the table, average accuracy on the 'operation' portion of the tasks was very high, no participant scored below 80% (the recommended cut-off to ensure that participants are performing both parts of the task).

TABLE 10 HERE

TABLE 11 HERE

**Predicting change detection from subitizing capacity.** We first examined whether perceptual capacity as measured by subitizing significantly predict change detection rates when controlling for working memory capacity using hierarchical regression. The first step included only the working memory span scores, subitizing range and estimation accuracy were added in the second step. The full regression is presented in Table 9. The first step of the regression including all the working memory span tasks did not significantly predict change detection (Adjusted $R^2$ = -.03, n.s.). The second step was significant (Adjusted $R^2$ = .18, $p < .001$) as predicted. As can be seen in Table 12 subitizing capacity and estimation accuracy were both significant predictors of change detection rates when controlling for working memory

TABLE 12 HERE

These results were fully replicated when the same analyses were run using detection sensitivity (A). The stepwise regression showed that subitizing significantly predicts detection sensitivity (Beta = .29, t = 2.26, p = .027) while controlling for working memory. In addition, as in Study 1 there was no relationship between response bias (b) in the change detection task and either subitizing (r(68) = -.05, n.s.) or estimation accuracy (r(68) = -.14, n.s.).

Thus in replication of Study 1, both subitizing capacity and estimation accuracy appear to measure distinct constructs, both of which are predictive of change detection performance. Study 3 further showed that this prediction cannot be explained by working memory or other general task-taking aptitudes as the prediction remains significant when controlling for complex span working memory capacity.

**Predicting MOT from subitizing capacity.** Next, we examined whether subitizing capacity can predict a more continuous measure of perceptual capacity as reflected in the MOT task while controlling for any shared variance with working memory capacity. We thus used a multiple regression in which the first step included working memory capacity as a control variable, and the second step included subitizing and estimation accuracy as in previous analyses. The full regression is presented in Table 13. The first step of the regression was significant (Adjusted $R^2$ = .09, p = .029). More importantly the second step was also significant (Adjusted $R^2$ = .22, p = .004), indicating common variance between subitizing and MOT capacity separate from any variance associated with working memory capacity, as predicted from our general perceptual capacity hypothesis and in line with load theory.

TABLE 13 HERE

FIGURE 10 HERE

Taken together, the findings that subitizing predicted both change detection and MOT independently of working memory supports the hypothesised perceptual capacity as the construct underlying common variance between subitizing and visual perception tasks. Moreover, the differences in the transient versus continuous nature of change blindness and MOT respectively, supports further the notion that their co-variance with subitizing reflects individual differences in a more generalized perceptual capacity.

**Testing the generalized perceptual capacity hypothesis.** The perceptual capacity hypothesis predicts that individual differences in MOT, Change detection and subitizing all depend upon the same underlying capacity, which is distinct from working memory capacity. To further examine this hypothesis it is necessary to establish the relationship between MOT and change detection. This relationship should not be attributable to working memory capacity but rather to perceptual capacity as previously measured with subitizing. Therefore MOT capacity should predict change detection performance when controlling for working memory span, but not when controlling for subitizing.

To test this hypothesis we first ran a regression that included the working memory span scores and estimation accuracy in its first step (thus controlling for executive working memory capacity and general cognitive factors involved in estimation but not perceptual capacity) and both working memory and tracking capacity (ENOT) in the second step. The findings (see Table 14) show that MOT significantly predicted change detection (Adjusted $R^2$ = .10, p = .04) when working memory was controlled in support of our first prediction.

TABLE 14 HERE

FIGURE 11 HERE

A further regression (Table 15) was run to test the second prediction that subitizing (rather than working memory) accounts for the common variance between MOT and change detection. This regression included working memory span measures, estimation accuracy and subitizing capacity as control variables in its first step and MOT capacity in the second step. The first step was significant (Adjusted $R^2$= .18, p = .003) indicating that subitizing and estimation significantly predicted change blindness as before, while the second step was not significant (Adjusted $R^2$ = .18, n.s.) indicating that MOT does not contribute unique variance to the prediction of change detection when controlling for subitizing in line with our hypothesis.

TABLE 15 HERE

Thus MOT capacity appears to significantly predict change detection when controlling for variance accounted for by working memory capacity, paralleling the profile observed with perceptual capacity for subitizing. Taken together the results indicate that a single underlying construct appears to underlie subitizing, change detection and MOT and this construct is distinct from working memory capacity. This is in accordance with our prediction of a general perceptual capacity as distinct from the capacity for cognitive control (as stipulated in load theory, e.g. Lavie et al. 2004).

**Reliability of measures.** The far larger number of trials in the enumeration task in Study 3 afforded a reliability analysis based on split half correlations (Note 2). The data for each task was split such that every-other trial throughout the task was assigned to one split, or the other, respectively. Spearman-brown corrected correlation coefficients are presented in Table 16. As can be seen in the table, positive correlations between the split halves were significant for each of the measures. Importantly, the reliability of estimation accuracy ($r = .89$) was higher than that of subitizing capacity ($r = .70$). Thus clearly the unique relationship between subitizing and the other perceptual measures is not attributed to greater reliability of the subitizing measure compared to the estimation measure.

TABLE 16 HERE

## Factor Analysis

**Principle components analysis.** The results reported thus far suggest that there are two distinct constructs underlying performance across the tasks, one representing perceptual capacity and another representing working memory capacity. To examine this possibility further we applied a factor analysis approach to the data.

Behavioural performance scores for each of the variables of interest (Subitizing, change detection sensitivity, MOT, OSPAN, RSPAN and SSPAN scores) were first entered into a principle components analysis (PCA) with an orthogonal (Varimax) factor rotation. The Keyser- Meyer-Olkin measure of sampling adequacy indicated that the sample was adequate (KMO = .59; Keiser, 1970; Field, 2009). Bartlett's test for sphericity was significant ($X^2(15) = 60.02$, $p < .001$) suggesting that there were sufficient inter-item correlations for PCA.

Principle components were extracted with Eigenvalues greater than 1 (Keiser, 1964). This resulted in a two-factor solution, and was supported also by examination of a scree plot in which there was a clear point of inflexion at the third factor. The first factor accounted for 35.25% of the overall variance and the second factor accounted for 22.44% (57.69% cumulatively).

The rotated factor loadings for each variable are presented in Table 1. As can be seen from the table, the first component is indicated by working memory variables, all of which have high, positive loadings. Conversely the second factor is indicated by the 'perceptual' variables which have similarly high and positive loadings. The results of this analysis therefore support the conclusion that two distinct and dissociable factors underlie working memory and perceptual capacity respectively. Interestingly, MOT appears to not just load on the perceptual component but also to moderately load on the first component of working memory. Thus in addition to perceptual resources, MOT appears to also involve some working memory capacity.

TABLE 17 HERE

FIGURE 12 HERE

**Confirmatory factor analysis.** As a further test of the hypothesis that two distinct factors underlie performance across the tasks we replicated the model provided by the PCA in a confirmatory factor analysis (CFA). This allowed us to formally compare the best fitting model provided by the PCA (above) to other possible models, providing a better insight into the nature of the latent structure of the data. This and other models were assessed using LISREL 8 (Scientific Software International, Inc.) using maximum likelihood estimation procedure based on the correlation matrix presented in Table 11.

First we tested the null hypothesis of an independence model (with no structure), and established that it was significantly different to the observed date, $\chi^2(15) = 67.517$, $p < .001$) thus a model with no structure does not fit the data well.

We then tested our hypothesised model which included two latent variables: One representing working memory, indicated by the three complex span tasks; the other representing perceptual capacity, indicated by subitizing, MOT and change detection sensitivity. In line with the PCA results, MOT was also allowed to cross-load on both latent factors.

This model fit the data well, the minimum fit function Chi-square test of difference to the observed data, was non-significant ($\chi^2(8)= 10.83$, $p = .21$) indicating that the model does not significantly differ from the observed data. The Standardised Root Mean Residual (SRMR) for the hypothesised model was .07, indicating a good fit (for which .08 or lower residual variance is required Tabachnick & Fidell, 2013). The Akaike Information Criterion (AIC) was 35.96; lower than both the independence and saturated models (79.52 and 42.0 respectively) suggesting that the fit was superior to these alternative models with either no relationships or with unstructured relationships between every variable. The comparative fit index (CFI) was .95, also indicating a good fit (Tabachnick and Fidell, 2013). Figure 13 represents the standardised factor loadings for the hypothesised model. All of the estimated factor loadings were significant ($p < .05$ for all). As can be seen in Figure 13, the loading of MOT on the 'perceptual' variable was numerically larger than that on the 'working memory' variable however both loadings were statistically significant.

FIGURE 13 HERE

Finally, we also tested a model representing the alternative hypothesis that all of the variance observed in the data may be best described by a single latent factor. The fit of this model was also poor and was rejected on the basis of a significant minimum fit function Chi-square test ($\chi^2(9) = 25.07$, $p = .003$). The AIC was 49.06, also higher than our hypothesized model demonstrating more unexplained variance.

## General Discussion

The present findings establish a common perceptual capacity limit for visual detection as measured in four different tasks: subitizing, MOT, load-induced blindness and in a change detection task often referred to as change blindness. Specifically, the findings show that an individual who is able to subitize a larger number of items from a brief display will have better accuracy and sensitivity for detection of changes in meaningful scenes (in the change detection task) as well as for detection of peripheral stimuli while attention is occupied in a central task (in the load-induced blindness paradigm).

Moreover, while subitizing necessitates rapid encoding from brief displays, the MOT task used in Study 3 involved continual tracking for several seconds and appears to rely upon the same underlying capacity. The results of Study 3 thus establish that the observed individual differences in perceptual capacity are common to continual deployment of attention to non-transient displays in the MOT task. Importantly, the results showed that subitizing capacity was consistently either a distinct factor alongside estimation ability; or the only unique factor after estimation ability was controlled for in multiple regression analyses. These results provide some evidence against general factors explaining the effect (for example motivation) since these would be reflected in the common variance in task performance both within and beyond the subitizing range.

In Study 3 in addition to the estimation accuracy we also explicitly controlled for working memory capacity using complex span tasks involving a variety of cognitive modalities. These tasks are specifically designed to measure working memory capacity in the face of distracting dual-task goals which load cognitive control resources. Complex span tasks have repeatedly been demonstrated as powerful predictors of various aspects of cognitive performance, including the control of attention (e.g. Redick & Engle, 2006) and general fluid intelligence (Redick et al. 2012; Foster et al. 2015; Unsworth et al. 2009). Furthermore, the influence of any particular domain-specific memory function was reduced by measuring the common variance between multiple tasks involving a variety of stimulus classes. Perceptual capacity as measured by subitizing (and object tracking) capacity was still a strong predictor of change detection when controlling for working memory capacity across these varied measures of working memory span.

Overall then the present results cannot be explained by a specific capacity for stimulus detection under transient, very brief presentations (since they extend to continuous tracking) nor can they be explained by a general cognitive ability or working memory capacity. Instead these findings support a construct of perceptual capacity underlying common variance between subitizing and visual perception tasks. The confirmatory factor analyses in Study 3 further supported this conclusion, showing that the best fitting model was one which included two latent factors which represented distinct perceptual capacity and working memory capacity. An alternative model which only included one ‘general’ latent factor did not fit the data well, supporting the hypothesis that subitizing and other tasks depend upon a distinct underlying attentional capacity for perceptual processing. Interestingly, in the best fitting model the capacity for tracking multiple moving objects loaded not only on the perceptual factor but also to a smaller extent on working memory. This is perhaps to be expected given that the only the MOT task involved extended durations (8 second trial) and required active maintenance of each target object despite the continuous change of each object positions throughout the full trial duration.

Of course each of the tasks used involved other specific sources of variance, for example, the ability to divide spatial attention between fixation and the periphery (in the load induced blindness task); a search (for a change) component (in the change detection task), perceptual grouping factors across motion (in the MOT task) or static stimuli (in the subitizing task) to name but a few. Such factors should account for some of the variance not explained by generalised perceptual capacity to perceive more items in parallel (e.g. irrespective of whether these are grouped or ungrouped). Despite the possible contribution of various different cognitive resources to task performance, the present results indicate that the perceptual capacity of attention is consistently a factor in tasks involving perceptual load. This therefore supports the importance of perceptual capacity as a key component of attentional processing across various task demands.

A common perceptual capacity limit of attention in tasks involving high perceptual load is consistent with previous studies of the load theory of attention. Previous research has provided support for a general capacity limit by demonstrating that various manipulations of load (e.g. feature versus conjunction discrimination in non-spatial search, increased set size in spatial search, object perception across rotation) converge upon the same result: reduced perception of stimuli outside the focus of attention. Here we find novel support from an individual differences perspective by establishing that an individual’s capacity limits are correlated across different visual perception tasks, and are distinct from capacity for higher level cognitive control.

The diversity of the tasks examined here further attests to generality of perceptual capacity limits. In the subitizing task numerical judgements were made based on a single, brief display of simple, square stimuli; whereas in the change detection task the visual display flickered repeatedly for several seconds, requiring a search among relatively complex real-world scenes; in the load-induced blindness task a line length discrimination task was combined with contrast increment detection in the periphery, and finally the MOT task involved a continuous display with minimal requirement for rapid encoding. Despite these differences, these tasks all recruit a common perceptual capacity.

The present results are consistent with a growing body of literature demonstrating that subitizing depends upon the allocation of attention so that subitizing capacity is reduced when subjects pay attention to another task, especially under conditions of high perceptual load (e.g. Vetter, Butterworth & Bahrami, 2008; 2011 and others, see General Introduction for review). Indeed the findings of a direct relation between subitizing capacity and the number of objects tracked in a motion tracking task (Chesney & Haladjian, 2012 and the present work) is highly suggestive of common perceptual capacity across tasks. Our findings complement this previous work in demonstrating common intra-individual perceptual capacity across subitizing, MOT and visual detection tasks while also controlling for non-perceptual variables. We note also a recent finding that the subitizing phenomenon typically assessed with simple shapes generalizes also to real-world stimuli (Railo, Karhu, Mast, Pesonen & Koivisto 2016). This is consistent with our demonstration that the capacity underlying subitizing generalizes to visual detection for both meaningful real-world stimuli and more elementary, simple shapes.

Our results establish the subitizing task as a simple quantification of an individual's general perceptual capacity limit that can predict their performance in visual detection tasks, and object tracking as a parallel measure of the same capacity for perception. As such they provide a potentially powerful indicator of individual abilities relevant to various tasks in industry, defence and security. Many roles depend on an individual's visual detection and object tracking capacity, for example x-ray screening and CCTV monitoring. The present research thus provides a scientific basis for devising future personnel selection tests for security and defence.

## Footnotes

1. Set size nine was excluded from analysis due to 'end effects' observed in previous research wherein participants tend to guess the maximum value when presented with large set sizes; artificially inflating the number of correct responses.
2. For the enumeration task in Studies 1-2 there were an insufficient number of trials to effectively fit the bilinear function with only half of the data.

## Acknowledgements

This work was supported in part by the Ministry of Defence, Defence Science and Technology Laboratory (DSTL). The authors would like to thank the London Science Museum for the opportunity and facilities for data collection and Mr Anders Jespersen, Ms Danielle Bagri and Ms Madeleine Payne for their assistance in data collection at the Museum.

## References

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological science*, *19*(4), 392-398.

Beck, D. M., Rees, G., Frith, C. D., & Lavie, N. (2001). Neural correlates of change detection and change blindness. *Nature neuroscience*, *4*(6), 645-650.

Boot, W. R., Brockmole, J. R., & Simons, D. J. (2005). Attention capture is modulated in dual-task situations. *Psychonomic Bulletin & Review*, *12*(4), 662-668.

Brand-D'Abrescia, M., & Lavie, N. (2008). Task coordination between and within sensory modalities: Effects on distraction. *Attention, Perception, & Psychophysics*, *70*(3), 508-515.

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision*, *10*(6), 20-20.

Carmel, D., Fairnie, J., & Lavie, N. (2012). Weight and see: loading working memory improves incidental identification of irrelevant faces. *Frontiers in psychology*, *3*, 286.

Carmel, D., Thorne, J. D., Rees, G., & Lavie, N. (2011). Perceptual load alters visual excitability. *Journal of experimental psychology: human perception and performance*, *37*(5), 1350.

Cartwright-Finch, U., & Lavie, N. (2007). The role of perceptual load in inattentional blindness. *Cognition*, *102*(3), 321-340.

Chesney, D. L., & Haladjian, H. H. (2011). Evidence for a shared mechanism used in multiple-object tracking and subitizing. *Attention, Perception, & Psychophysics*, *73*(8), 2457-2480.

De Fockert, J. W. (2013). Beyond perceptual load and dilution: a review of the role of working memory in selective attention. *Frontiers in psychology*, *4*.

De Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science*, *291*(5509), 1803-1806.

Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PloS one*, *10*(4), e0121945.

Egeth, H. E., Leonard, C. J., & Palomares, M. (2008). The role of attention in subitizing: Is the magical number 1?. *Visual Cognition*, *16*(4), 463-473.

Forster, S., & Lavie, N. (2008). Attentional capture by entirely irrelevant distractors. *Visual Cognition*, *16*(2-3), 200-214.

Forster, S., & Lavie, N. (2008). Failures to ignore entirely irrelevant distractors: the role of load. *Journal of Experimental Psychology: Applied*, *14*(1), 73.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & cognition*, *43*(2), 226-236.

Hittner, J. B., May, K., & Silver, N. C. (2003). A Monte Carlo evaluation of tests for comparing dependent correlations. *The Journal of general psychology*, *130*(2), 149-168.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *The American journal of psychology*, *62*(4), 498-525.

Konstantinou, N., & Lavie, N. (2013). Dissociable roles of different types of working memory load in visual detection. *Journal of experimental psychology: human perception and performance*, *39*(4), 919.

Lavie, N. (2000). Selective attention and cognitive control: Dissociating attentional functions through different types of load. *Attention and performance XVIII*, 175-194.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in cognitive sciences*, *9*(2), 75-82.

Lavie, N. (2006). The role of perceptual load in visual awareness. *Brain research*, *1080*(1), 91-100.

Lavie, N., Beck, D. M., & Konstantinou, N. (2014). Blinded by the load: attention, awareness and the role of perceptual load. *Phil. Trans. R. Soc. B*, *369*(1641), 20130205.

Lavie, N., & Cox, S. (1997). On the efficiency of visual selective attention: Efficient visual search leads to inefficient distractor rejection. *Psychological Science*, *8*(5), 395-396.

Lavie, N., & Fox, E. (2000). The role of perceptual load in negative priming. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(3), 1038.

Lavie, N., & Torralbo, A. (2010). Dilution: A theoretical burden or just load? A reply to Tsal and Benoni (2010).

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339.

Macdonald, J. S., & Lavie, N. (2008). Load induced blindness. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1078.

Mack, A., & Rock, I. (1998). *Inattentional blindness* (Vol. 33). Cambridge, MA: MIT press.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive psychology*, *7*(4), 480-494.

Neisser, U. (1979). The concept of intelligence. *Intelligence*, *3*(3), 217-227.

Olivers, C. N., & Watson, D. G. (2008). Subitizing requires attention. *Visual Cognition*, *16*(4), 439-462.

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior research methods*, *47*(4), 1343-1355.

Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*, *121*(1), 147-153.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Railo, H., Karhu, V. M., Mast, J., Pesonen, H., & Koivisto, M. (2016). Rapid and accurate processing of multiple objects in briefly presented scenes. *Journal of vision*, *16*(3), 8-8.

Railo, H., Koivisto, M., Revonsuo, A., & Hannula, M. M. (2008). The role of attention in subitizing. *Cognition*, *107*(1), 82-104.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink?. *Journal of experimental psychology: Human perception and performance*, *18*(3), 849.

Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology*, *20*(5), 713-721.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*.

Remington, A., Cartwright-Finch, U., & Lavie, N. (2014). I can see clearly now: the effects of age and perceptual load on inattentional blindness.

Rensink, R. A. (2002). Change detection. *Annual review of psychology*, *53*(1), 245-277.

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., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*, *80*(1), 159-177.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059-1074.

Tabachnick, B. G., and Fidell, L. S. (2013). *Using multivariate statistics, 6th ed.* Boston: Pearson.

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(2), 331.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological review, 101*(1), 80.

Tuholski, S. W., Engle, R. W., & Baylis, G. C. (2001). Individual differences in working memory capacity and enumeration. *Memory & cognition*, *29*(3), 484-492.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, *17*(6), 635-654.

Vetter, P., Butterworth, B., & Bahrami, B. (2008). Modulating attentional load affects numerosity estimation: evidence against a pre-attentive subitizing mechanism. *PLoS One, 3*(9), e3269.

Vetter, P., Butterworth, B., & Bahrami, B. (2011). A candidate for the attentional bottleneck: set-size specific modulation of the right TPJ during attentive enumeration. *Journal of Cognitive Neuroscience*, *23*(3), 728-736.

Ward, E. J., & Scholl, B. J. (2015). Inattentional blindness reflects limitations on perception, not memory: Evidence from repeated failures of awareness. *Psychonomic Bulletin & Review*, *22*(3), 722-727.

Ward, E. J., Bear, A., & Scholl, B. J. (2016). Can you perceive ensembles without perceiving individuals?: The role of statistical perception in determining whether awareness overflows access. *Cognition*, *152*, 78-86.

Xu, X., & Liu, C. (2008). Can subitizing survive the attentional blink? An ERP study. *Neuroscience letters*, *440*(2), 140-144.

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika*, *70*(1), 203-212.

# Tables

Table 1

Average change detection performance.

| Measure | Mean (SD) |
|---|---|
| Detection Rate | .68 (.13) |
| False Alarm Rate | .44 (.06) |
| A | .89 (.06) |
| b | 1.99 (0.49) |

Table 2

Study 1: Correlation Matrix.

| | 1 | 2 | 3 |
|---|---|---|---|
| 1.Subitizing | - | | |
| 2.Estimation | .24** (.14 - .34) | - | |
| 3. Detection Sensitivity (A) | .37** (.27 - .48) | .15* (.06 - .25) | - |
| 4. Detection Rate | .31** (.20 - .42) | .15* (.04 - .25) | .80** (.75 - .84) |

Note: Detection sensitivity and detection rate refer to performance on the change detection 'flicker' task. * indicates $p < .05$, ** indicates $p < .005$. Upper and lower 95% Confidence intervals are presented in brackets.

Table 3

Hierarchical regression predicting change detection rates from subitizing (step 2) while controlling for estimation accuracy (step 1).

| Model | Predictor | Beta | t | p |
|---|---|---|---|---|
| 1 | Constant | | 37.79 | <.001 |
| Adjusted $R^2$= .02 | Estimation | .24 | 2.49 | .013 |
| p = .013 | | | | |
| 2 | Constant | | 11.57 | <.001 |
| Adjusted $R^2$= .14 | Estimation | .24 | 4.12 | <.001 |
| p < .001 | Subitizing | .36 | 6.35 | <.001 |

Table 4

Hierarchical regression predicting change detection sensitivity from subitizing (step 2) while controlling for estimation accuracy (step 1).

| Model | Predictor | Beta | t | p |
|---|---|---|---|---|
| 1 | Constant | | 114.75 | <.001 |
| Adjusted $R^2$= .02 | Estimation | .15 | 2.50 | .013 |
| p = .013 | | | | |
| Adjusted $R^2$= .19 | Constant | | 47.16 | <.001 |
| p < .001 | Estimation | .25 | 4.55 | <.001 |
| | Subitizing | .43 | 7.74 | <.001 |

Table 5

Load-induced blindness task performance measures.

| Measure | Mean (SD) |
|---|---|
| Detection Rate | .70 (.21) |
| False Alarm Rate | .24 (.18) |
| A | .80 (.12) |
| B | 1.16 (0.55) |

Table 6

Study 2: Correlation matrix

| | 1 | 2 | 3 |
|---|---|---|---|
| 1.Subitizing | - | | |
| 2.Estimation | .25** (.08 - .44) | - | |
| 3. Detection Sensitivity (A) | .38** (.22 - .52) | .25** (.08 - .41) | - |
| 4. Detection Rate | .29** (.12 - .43) | .22* (.05 - .38) | .69** (.58 - .78) |

Note: Detection sensitivity and detection rate refer to performance on the load-induced blindness task. A single * indicates $p < .05$, a double ** indicates $p < .005$. 95% Confidence intervals are presented in brackets.

Table 7

Hierarchical regression predicting detection rate in the load-induced blindness task from subitizing capacity (step 2) while controlling for estimation accuracy (step 1).

| Model | Predictor | Beta | t | p |
|---|---|---|---|---|
| 1 | Constant | | 5.48 | <.001 |
| Adjusted $R^2$= .04 | Estimation | .22 | 2.51 | .014 |
| P = .014 | | | | |
| Adjusted $R^2$= .09 | Constant | | 3.43 | .001 |
| P = .008 | Estimation | .16 | 1.75 | .083 |
| | Subitizing | .24 | 2.70 | .008 |

Table 8

Hierarchical regression predicting detection sensitivity in the load-induced blindness task from subitizing capacity (step 2) while controlling for estimation accuracy (step 1).

| Model | Predictor | Beta | t | p |
|---|---|---|---|---|
| 1 | Constant | | 13.71 | <.001 |
| Adjusted $R^2$= .05 | Estimation | .25 | 2.77 | .006 |
| p = .006 | | | | |
| Adjusted $R^2$= .15 | Constant | | 10.43 | <.001 |
| p < .001 | Estimation | .16 | 1.77 | .080 |
| | Subitizing | .34 | 3.89 | <.001 |

Table 9

Change detection performance in Study 3.

| Measure | Mean (SD) |
|---|---|
| Detection rate | .50 (.12) |
| False Alarm Rate | .01 (.09) |
| Detection Sensitivity (A) | .82 (.08) |
| Bias (b) | 2.41 (0.69) |

Table 10

Average SPAN score and operation accuracy in each of the complex span tasks.

| | Mean Score (SD) | Mean Accuracy (SD) |
|---|---|---|
| OSPAN | 25.27 (5.34) | 96.74 (3.48) |
| RSPAN | 23.50 (4.51) | 95.81 (3.90) |
| SSPAN | 18.09 (4.62) | 98.04 (2.65) |

Table 11

Study 3: Correlation Matrix

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1.Subitizing | - | | | | | | |
| 2.Estimation | .12 | - | | | | | |
| | (-.11, .35) | | | | | | |
| 3.Detection | .27* | .19 | - | | | | |
| sensitivity (A) | (.07, .45) | (-.09, .42) | | | | | |
| 4.Detection | .39** | .33* | .57** | - | | | |
| rate | (.17, .57) | (.08, .53) | (.40, .73) | | | | |
| 5.MOT | .37** | .32* | .30* | .30* | - | | |
| (ENOT) | (.18, .53) | (.13, .50) | (.07, .51) | (.09, .50) | | | |
| 6.OSPAN | -.11 | .33* | .06 | .03 | .27* | - | |
| | (.40, .20) | (.12, .50) | (-.24, .33) | (-.18, .20) | (.07, .46) | | |
| 7.RSPAN | .26* | .29* | .04 | .08 | .29* | .44** | - |
| | (.04, .48) | (.09, .47) | (-.23, .30) | (-.16, .29) | (.06, .49) | (.20, .60) | |
| 8.SSPAN | .05 | .40** | .06 | .11 | .26* | .35** | .28** |
| | (-.20, .31) | (.21, .57) | (-.15, .27) | (-.12, .32) | (.04, .45) | (.11, .56) | (.09, .47) |

Note: Detection sensitivity and detection rate refer to performance on the change detection ‘flicker’ task. * indicates p < .05, a double ** indicates p < .005. Upper and lower 95% Confidence intervals are presented in brackets.

Table 12

Hierarchical regression predicting change detection from complex span working memory capacity (step 1), subitizing and estimation accuracy (step 2).

| Model | Variable | Beta | t | p |
|---|---|---|---|---|
| 1 | Constant | | 4.585 | <.001 |
| Adjusted $R^2 = -.03$, | OSPAN | -.032 | -.224 | .823 |
| n.s. | RSPAN | .059 | .424 | .673 |
| | SSPAN | .108 | .806 | .423 |
| 2 | Constant | | 2.641 | .010 |
| Adjusted $R^2 = .18$, | OSPAN | .028 | .206 | .837 |
| $p < .001$ | RSPAN | -.126 | -.950 | .346 |
| | SSPAN | -.006 | -.048 | .962 |
| | Estimation | .310 | 2.469 | .016 |
| | Subitizing | .390 | 3.270 | .002 |

Table 13

Hierarchical regression predicting MOT capacity from working memory capacity (step 1), subitizing, and estimation accuracy (step 2).

| Model | Variable | Beta | t | p |
|---|---|---|---|---|
| 1 | Constant | | .085 | .933 |
| Adjusted $R^2$ = .09, | OSPAN | .134 | .996 | .323 |
| p = .029 | RSPAN | .187 | 1.429 | .158 |
| | SSPAN | .156 | 1.240 | .219 |
| 2 | Constant | | -1.559 | .124 |
| Adjusted $R^2$ = .22, | OSPAN | .210 | 1.599 | .115 |
| p = .004 | RSPAN | .032 | .247 | .806 |
| | SSPAN | .090 | .732 | .467 |
| | Estimation | .160 | 1.303 | .198 |
| | Subitizing | .361 | 3.088 | .003 |

Table 14

Hierarchical regression predicting change detection rate from working memory capacity and estimation accuracy (step 1) followed by MOT capacity (step 2).

| Model | Variable | Beta | t | Sig. |
|---|---|---|---|---|
| 1 | Constant | | 4.382 | <.001 |
| Adjusted $R^2$= | OSPAN | -.090 | -.651 | .518 |
| .06, n.s. | RSPAN | .015 | .113 | .910 |
| | SSPAN | -.002 | -.011 | .991 |
| | Estimation | .352 | 2.628 | .011 |
| 2 | Constant | | 4.518 | <.001 |
| Adjusted $R^2$ = | OSPAN | -.116 | -.852 | .398 |
| .10, p = .041 | RSPAN | -.026 | -.196 | .845 |
| | SSPAN | -.025 | -.193 | .848 |
| | Estimation | .301 | 2.262 | .027 |
| | MOT | .254 | 2.007 | .041 |

Table 15

Hierarchical regression predicting change detection rate from working memory capacity, estimation accuracy and subitizing (step 1) followed by MOT capacity (step 2).

| Model | Variable | Beta | t | p |
|---|---|---|---|---|
| 1 | Constant | | 2.641 | .010 |
| Adjusted $R^2$ = | OSPAN | .028 | .206 | .837 |
| .18, p = .003 | RSPAN | -.126 | -.950 | .346 |
| | SSPAN | -.006 | -.048 | .962 |
| | Estimation | .310 | 2.469 | .016 |
| | Subitizing | .390 | 3.270 | .002 |
| 2 | Constant | | 2.779 | .007 |
| Adjusted $R^2$ = | OSPAN | .001 | .007 | .994 |
| .18, n.s. | RSPAN | -.130 | -.980 | .331 |
| | SSPAN | -.017 | -.138 | .890 |
| | Estimation | .289 | 2.275 | .026 |
| | Subitizing | .344 | 2.686 | .009 |
| | MOT | .127 | .979 | .332 |

Table 16

Split-half correlations for each task, all were significant and positive ($p < .001$ for all).

| Measure | Spearman-Brown Corrected Correlation Coefficient |
|---|---|
| Change Detection (Study 1) | .69 |
| Change Detection (Study 3) | .72 |
| Load-Induced Blindness | .67 |
| Subitizing | .70 |
| Estimation | .89 |
| MOT | .87 |
| OSPAN | .95 |
| RSPAN | .84 |
| SSPAN | .91 |

Table 17

Varimax rotated factor loadings of each behavioural variable on both factors produced by the PCA.

| Measure | Component 1 Loading | Component 2 Loading |
|---|---|---|
| OSPAN | .83 | -.09 |
| RSPAN | .70 | .23 |
| SSPAN | .69 | .05 |
| Subitizing | -.01 | .81 |
| Change Detection (A) | -.01 | .69 |
| MOT | .42 | .65 |

Figures

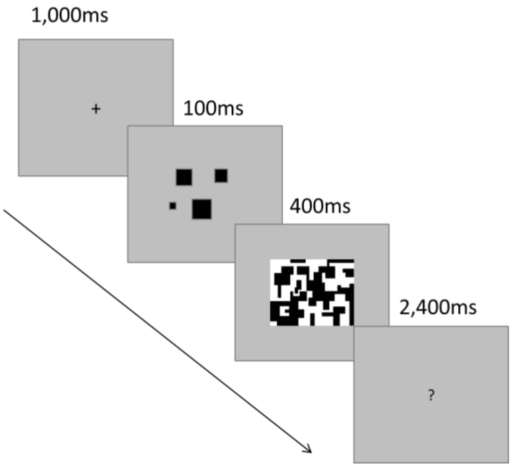


*Figure 1.* The enumeration task used in both Studies 1 and 2. A fixation cross was presented for 1 s which was then replaced by a display containing a variable number of (1-9) randomly sized and positioned squares. After 100 ms the stimulus display was replaced by a black and white noise mask for 400 ms and then a central ‘?’ for 2,400 ms. Participants were instructed to respond as quickly as possible after the squares were presented indicating how many they thought there were by pressing the corresponding key on the keyboard number pad.

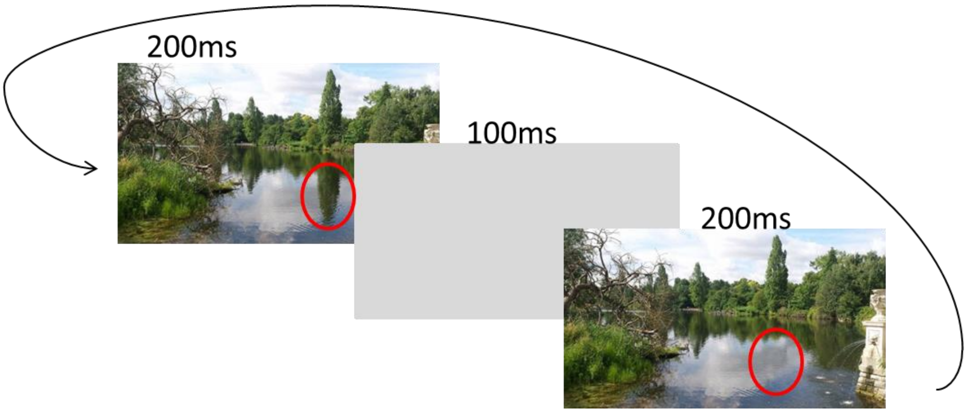


*Figure 2.* The change detection task used in Study 1. A photographic image of a real-world scene was presented for 200 ms, followed immediately by a grey rectangle of matching size and position and then another scene image for 200 ms, creating the appearance of a 'flickering' image. This stimulus cycle repeated for up to 15 s until the participant responded by pressing either the left or right 'shift' key to indicate whether both images were identical or not, respectively.

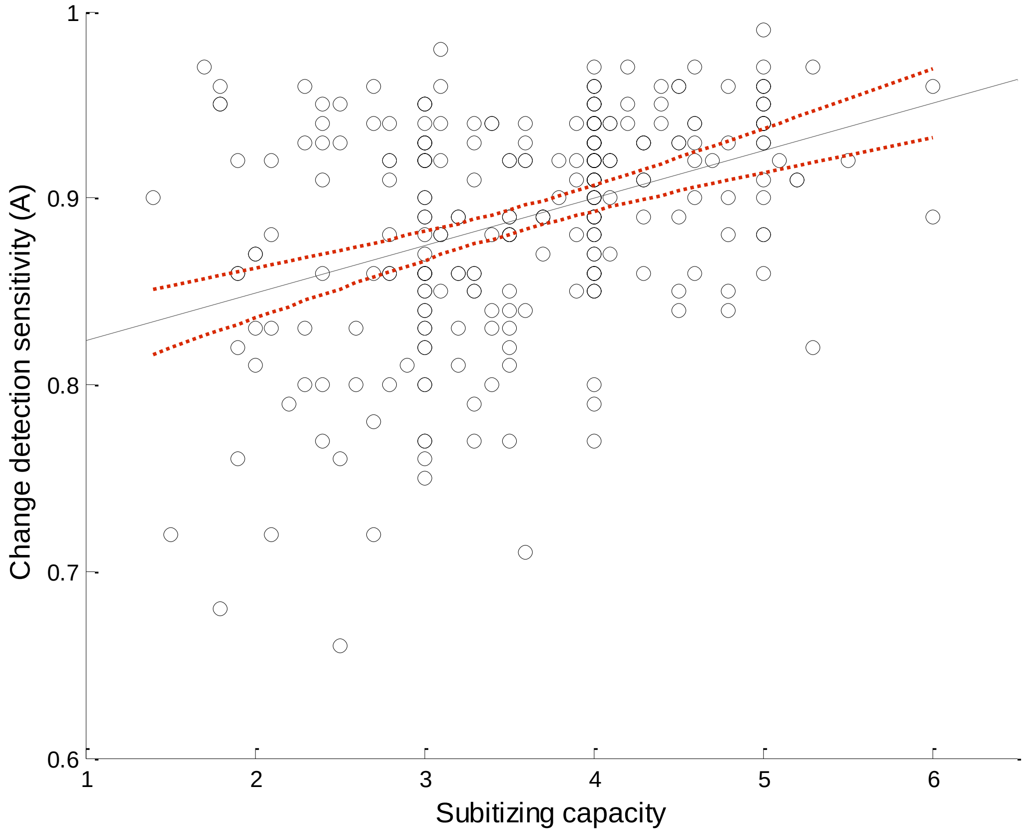


*Figure 3.* Change detection sensitivity and subitizing capacity correlation in Study 1. Dashed lines indicate 95% confidence intervals for the correlation.

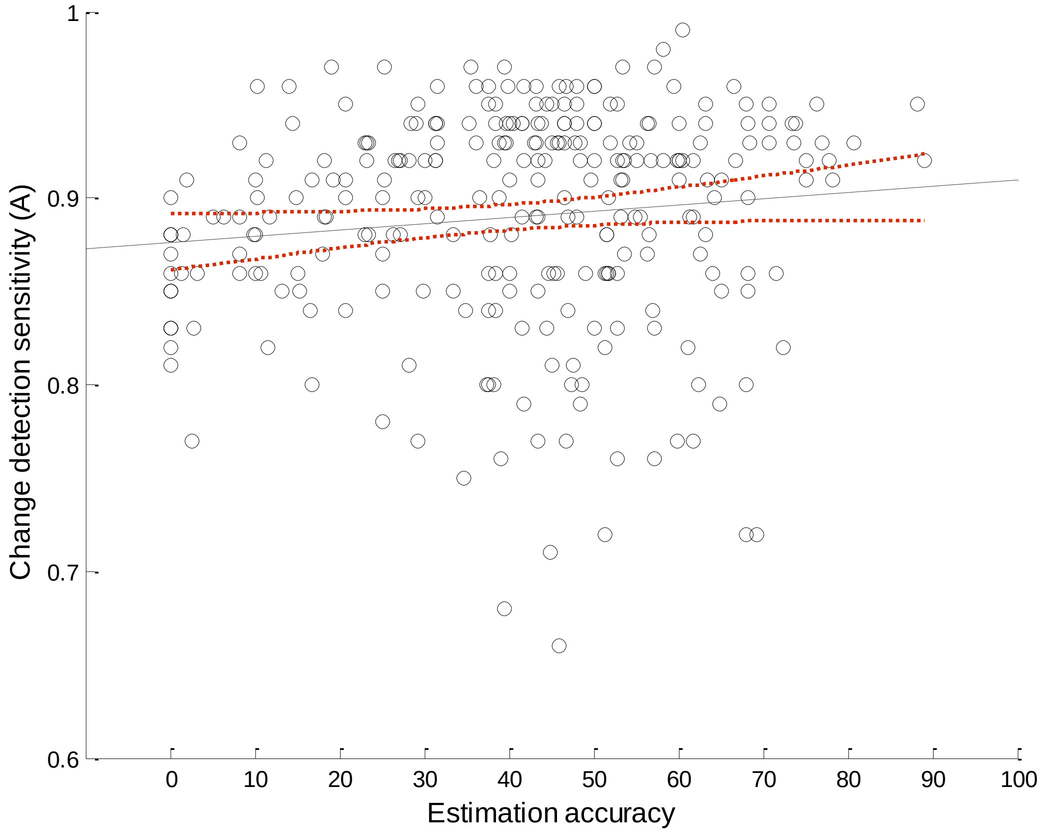


*Figure 4.* Change detection sensitivity and estimation accuracy correlation in Study 1. Dashed lines indicate 95% confidence intervals for the correlation.

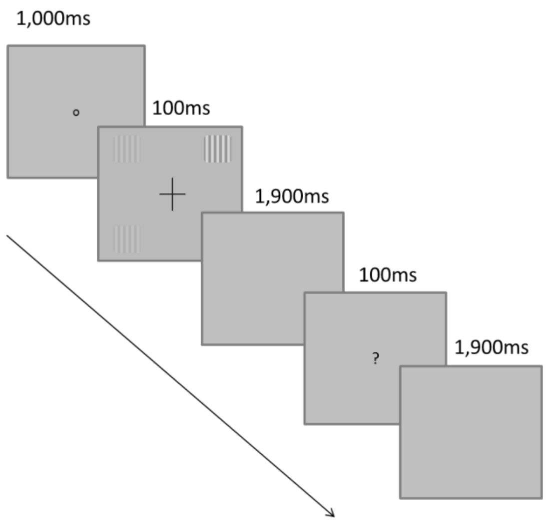


*Figure 5.* Example trial for the load-induced blindness task used in Study 2. Participants responded to the cross task during the first blank interval and then upon the presentation of the question mark symbol responded to the detection task.

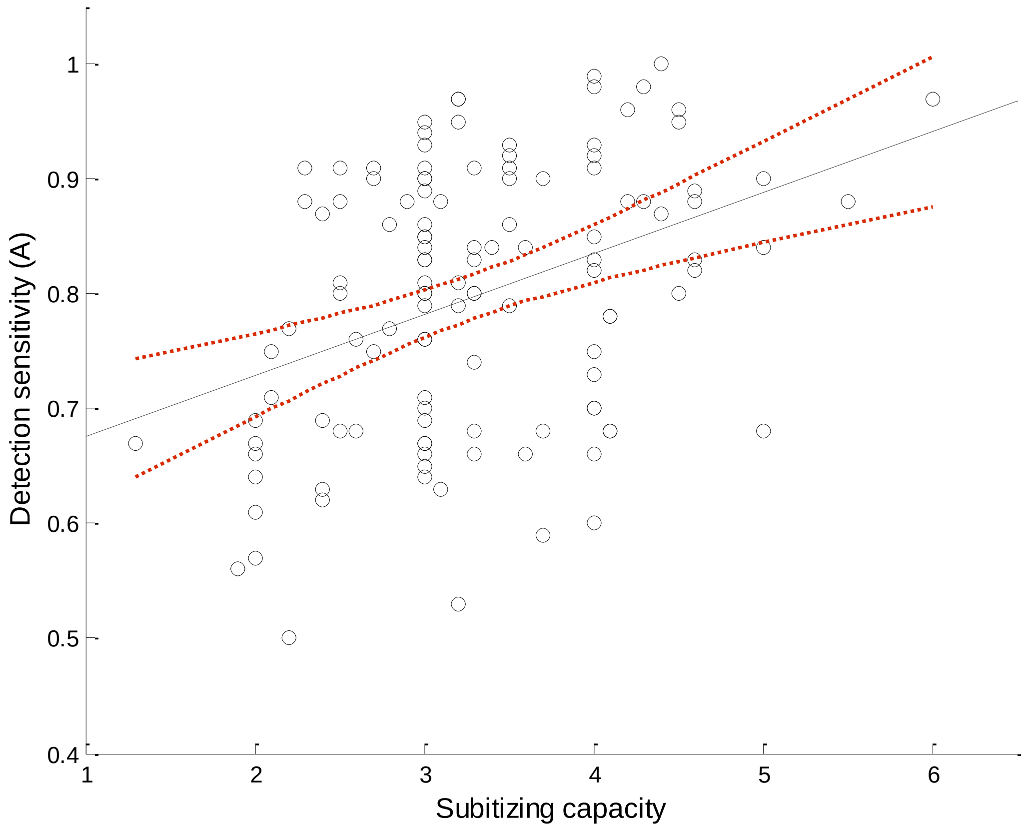


*Figure 6.* Detection sensitivity in the load-induced blindness task and subitizing capacity in Study 2. Dashed lines represent 95% confidence intervals for the correlation.

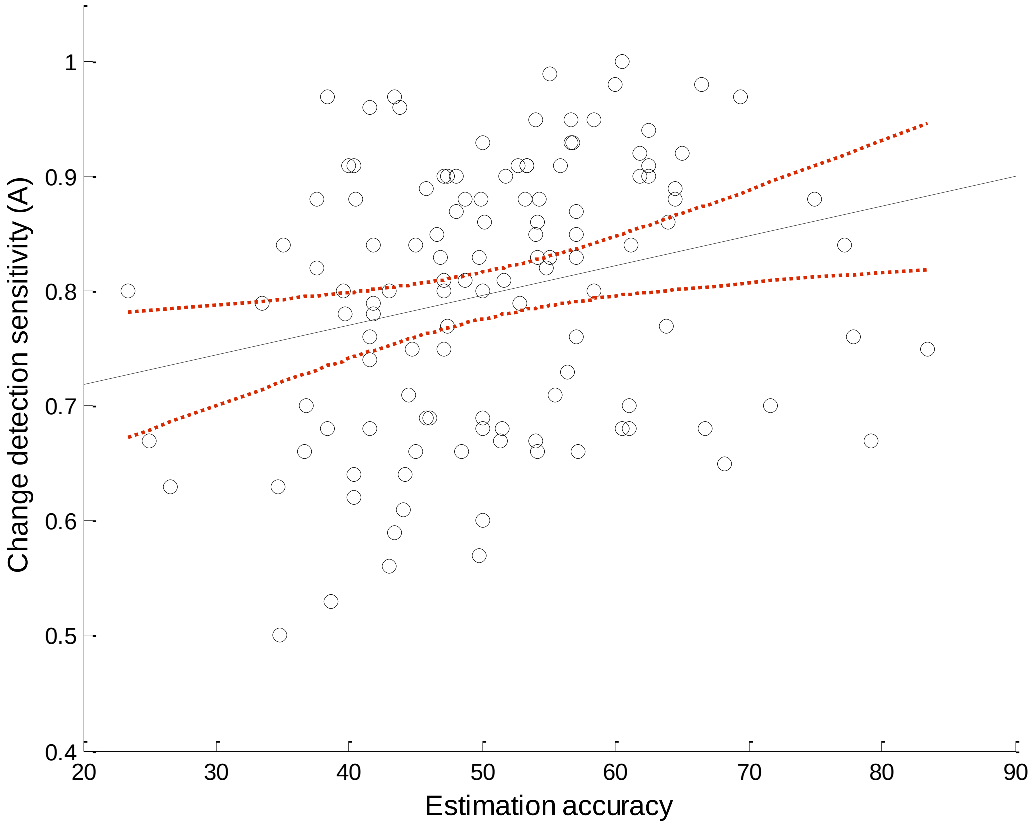


*Figure 7.* Detection sensitivity in the load-induced blindness task and estimation accuracy in Study 2. Dashed lines represent 95% confidence intervals for the correlation.

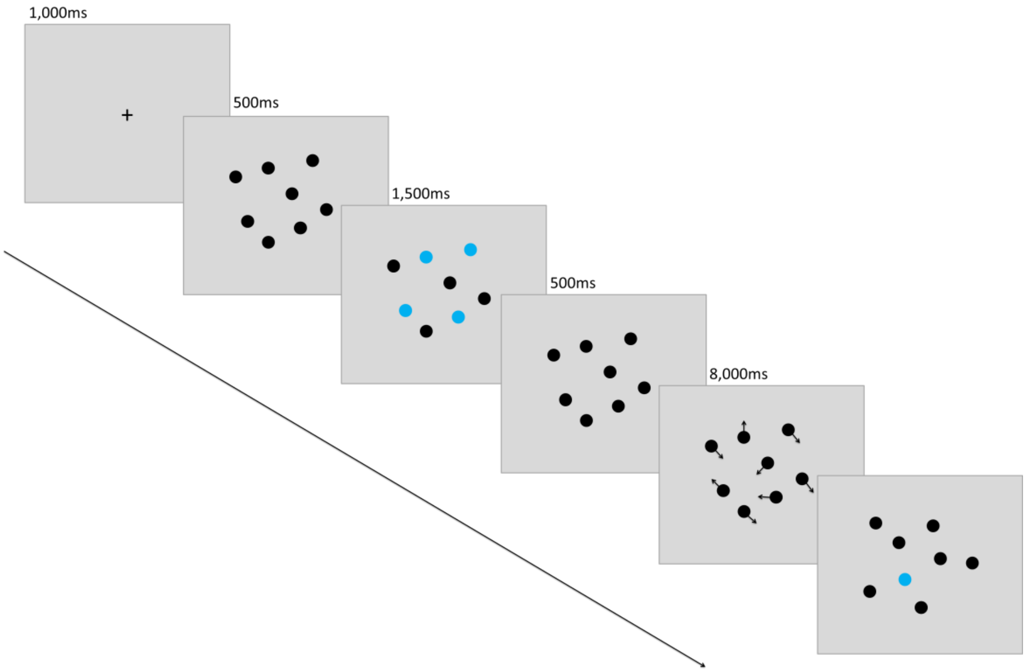


*Figure 8.* The MOT task used in Study 3. Participants were presented with eight black dots, four of which briefly turned blue before turning back to black, all eight dots then moved around randomly for 8 seconds. The dots then stopped moving, one turned blue again and the participant responded by indicating whether or not this was one of the original targets.

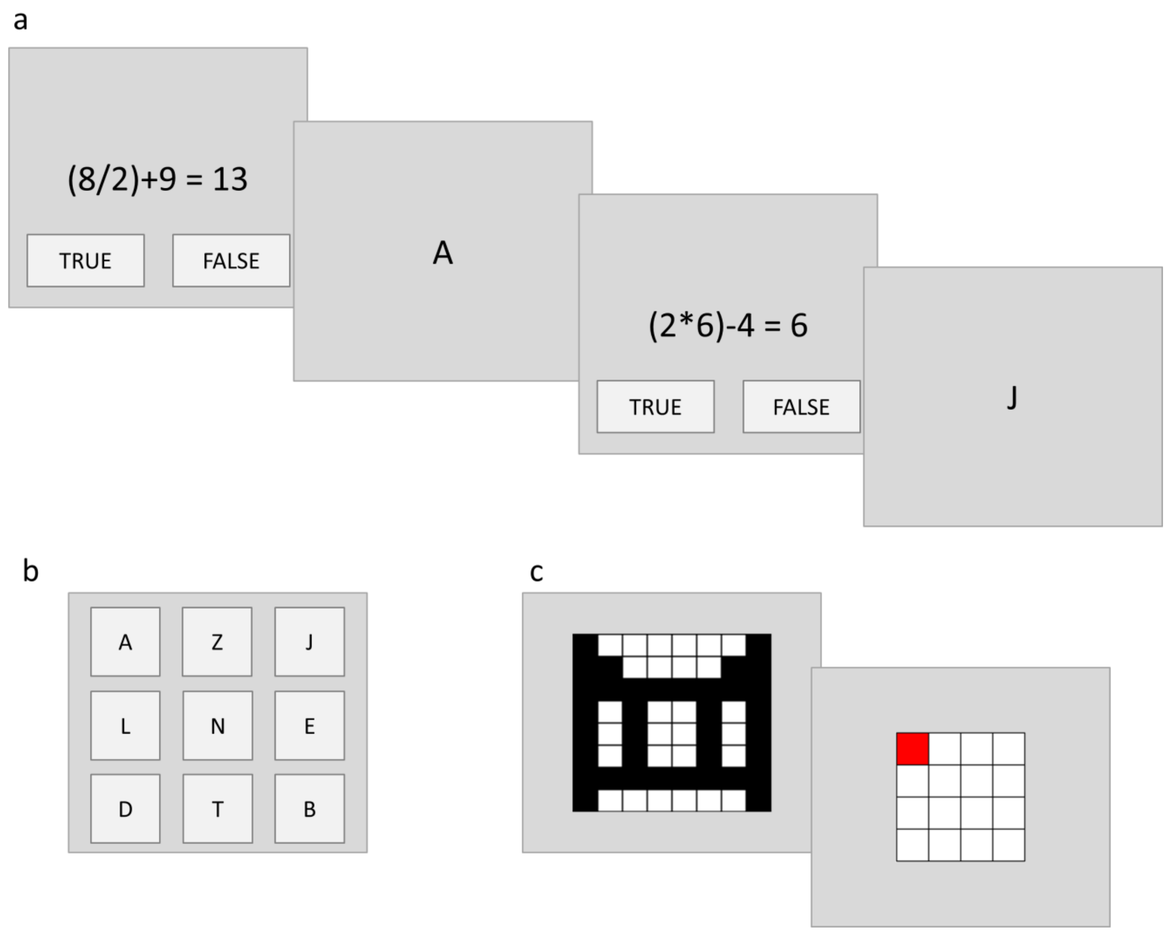


*Figure 9.* A typical trial in the OSPAN task (a). Participants perform a series of sums and memorise subsequently displayed letters. After a variable number of trials a memory test screen (b) is presented. In the SSPAN task symmetry judgements take the place of sums and location probes take the place of letter memoranda (c).

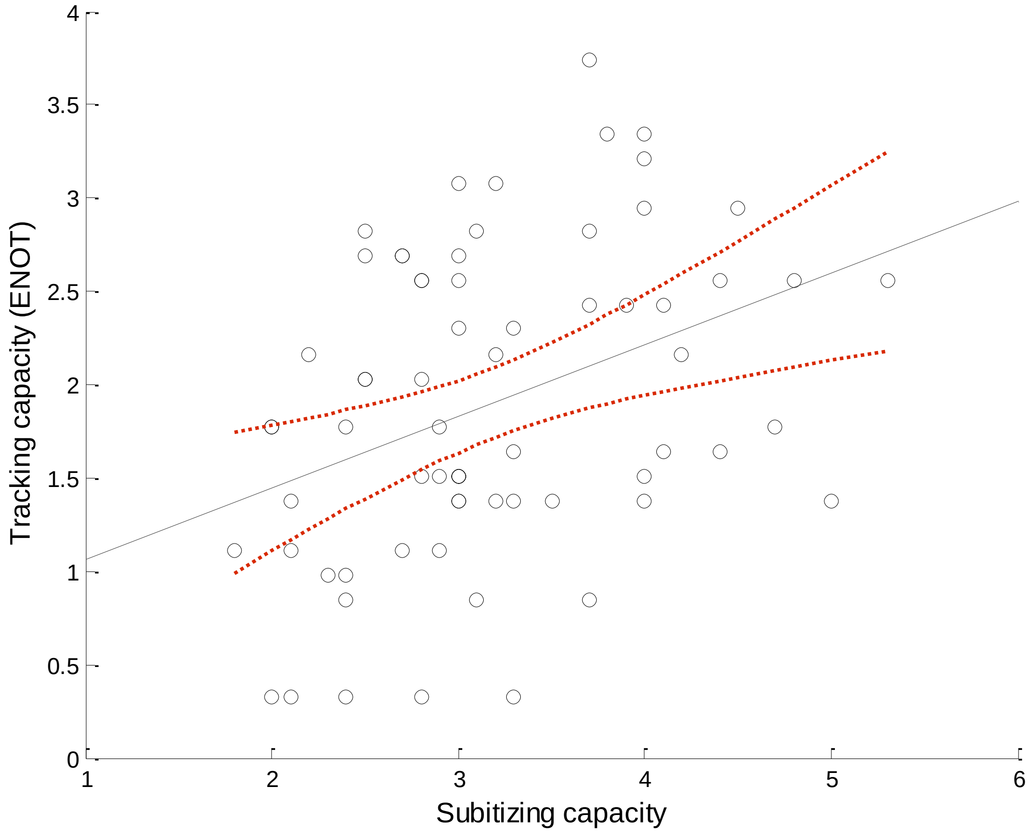


*Figure 10.* Object tracking capacity (ENOT) and subitizing capacity correlation in Study 3. Dashed lines represent 95% confidence intervals for the correlation.

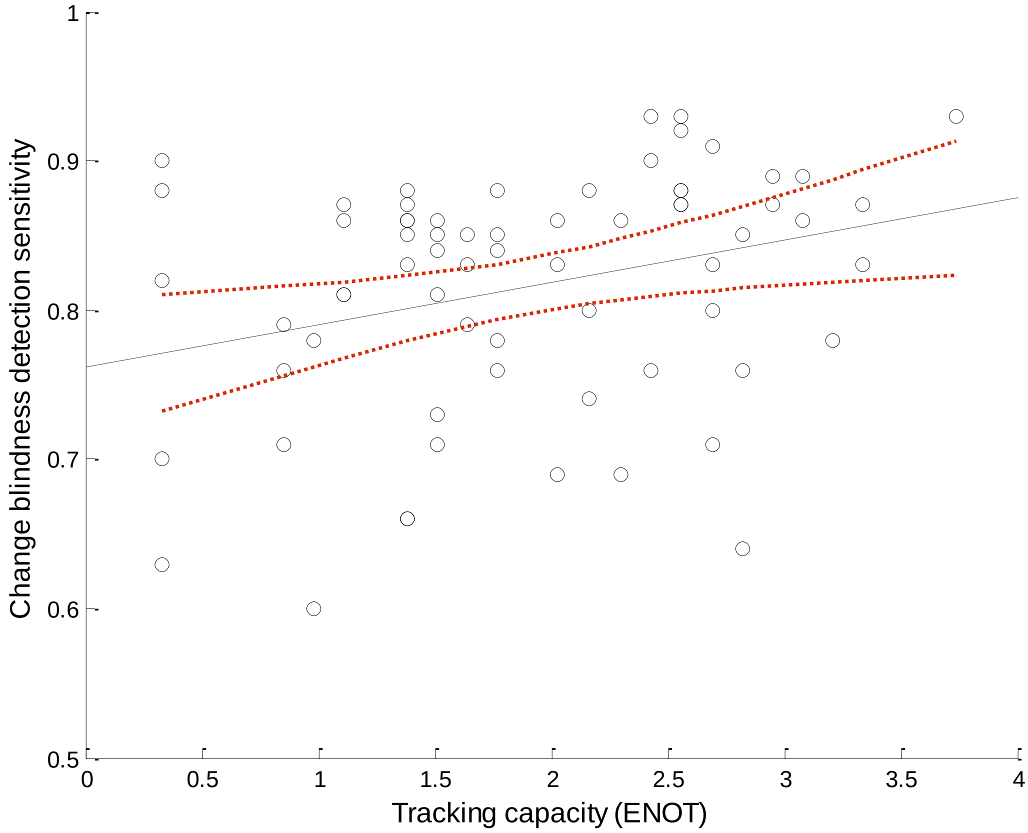


*Figure 11*. Object tracking capacity (ENOT) and change detection rate correlation in Study 3. Dashed lines represent 95% confidence intervals for the correlation.

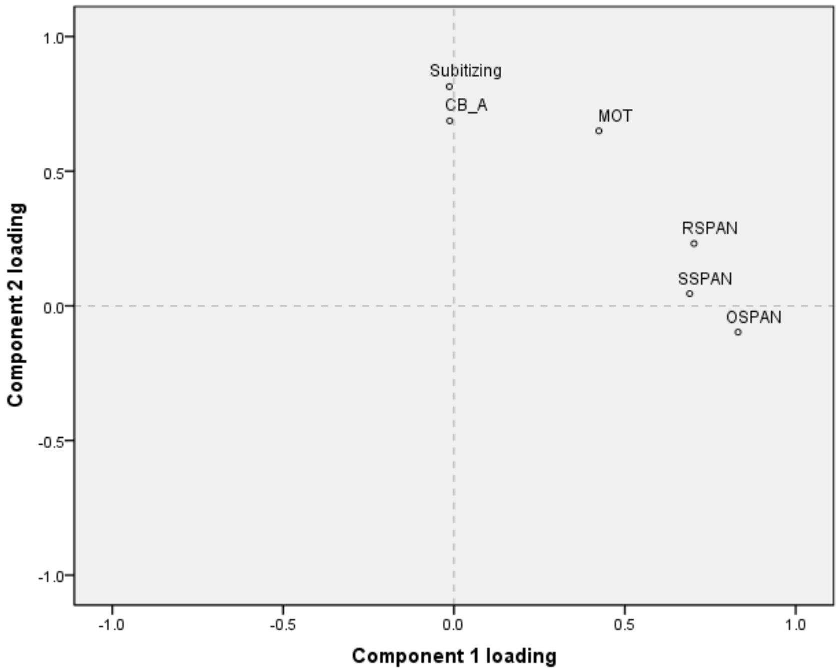


*Figure 12.* Factor loadings of each variable: subitizing capacity (Subitizing), change detection sensitivity (CB_A) MOT, RSPAN, SSPAN and OSPAN in Varimax-rotated space. Note that dashed lines represent a loading of zero on each component axis.

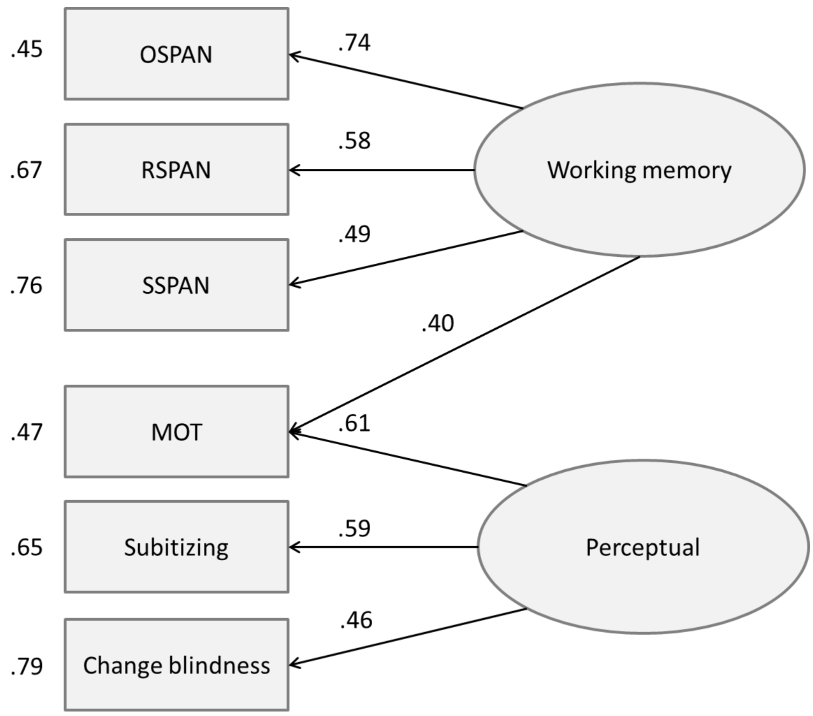


*Figure 13.* Confirmatory Factor Analysis results. Residuals (shown on the left of the variables) and factor loadings are based on the ‘completely standardised solution’ from LISREL 8.