Social prioritisation in scene viewing and the effects of a spatial memory load

A. P. Martinez-Cedillo[1] · Kevin Dent[2] · Tom Foulsham[2]

Abstract

When free-viewing scenes, participants tend to preferentially fixate social elements (e.g., people). In the present study, we tested whether this bias would be disrupted by increasing the demands of a secondary dual-task: holding a set of (one or six) spatial locations in memory, presented either simultaneously or sequentially. Following a retention interval, participants judged whether a test location was present in the to-be-remembered stimuli. During the retention interval participants free-viewed scenes containing a social element (a person) and a non-social element (an object) that served as regions of interest. In order to assess the impact of physical salience, the non-social element was presented in both an unaltered baseline version, and in a version where its salience was artificially increased. The results showed that the preference to look at social elements decreased when the demands of the spatial memory task were increased from one to six locations, regardless of presentation mode (simultaneous or sequential). The high-load condition also resulted in more central fixations and reduced exploration of the scene. The results indicate that the social prioritisation effect, and scene viewing more generally, can be affected by a concurrent memory load.

Introduction

The world in which human behaviour takes place is complex and cluttered, with many possible stimuli to analyse and act upon. According to current conceptualisations of attention, stimuli compete for access to neural and computational resources, which in turn govern access to behavioural responses (e.g., Desimone & Duncan, 1995). The question of how and why particular stimuli may be prioritised during this competition is an old one (e.g., see Moray, 1959), which has often focussed on the relative importance of physical stimulus properties, in comparison to the meaning that the stimuli hold for the observer.

Recently this debate has played out in the context of the question of which stimuli "capture attention" when we view scenes. According to one tradition, physical properties such as "saliency from feature contrast" are particularly important (see Theeuwes, 2010, for a review). For instance, high physical salience distractors tend to be targets for early saccades, even when the goal is to explicitly ignore them (Anderson & Donk, 2017; Donk & Van Zoest, 2008; Foulsham & Underwood, 2008; Underwood & Foulsham, 2006). Other research has emphasised the importance of the meaning of objects for predicting saccades. In particular, even when objects are of similar physical salience, participants tend to preferentially direct fixations to items of social significance (e.g., people; see Birmingham et al., 2008a, b, 2009; End & Gamer, 2017, 2019).

Recent research has begun to investigate the behavioural characteristics of this prioritisation of social objects. We recently investigated whether the social prioritisation effect was reduced when verbal working memory resources were taxed (Martinez-Cedillo, Dent & Foulsham, 2022). Increasing verbal working memory load is a technique known to disrupt goal-based control of attention in the face of a salient distractor (e.g., Boot et al., 2005; Burnham, 2010; Lavie & De Fockert, 2005). Our results showed that the social prioritisation effect was undiminished under the conditions of

✉ A. P. Martinez-Cedillo
   priscilla.martinezcedillo@york.ac.uk

1   Department of Psychology, University of York,
    York YO10 5DD, UK

2   Department of Psychology, University of Essex, Wivenhoe
    Park, Colchester, Essex CO4 3SQ, UK

a high memory load, demonstrating the robust and perhaps obligatory nature of this effect.

Bianchi et al. (2020) also recently investigated the effects of a cognitive load on looking behaviour in real and depicted environments in which a confederate was present. Their results showed that looking towards the confederate was disrupted when participants were given a cognitive load. However, in this study the cognitive load task took the form of an n-back task in which participants were continuously presented with distracting information. Given the close links between auditory and visual spatial attention (e.g., see Spence & Driver, 1996), the disruption involved could well be occurring not because of a domain-general cognitive load, but because of specific interference with visual spatial attention. Furthermore, in the study of Bianchi et al. (2020), a cognitive load condition is compared against a condition involving no load, meaning that a high cognitive load is confounded with other more general dual-task scheduling processes. In contrast, the study by Martinez-Cedillo et al. (2022) compares low and high cognitive load conditions, meaning that only the amount of load differs between conditions, and other more general dual-task processes are held constant. In addition, in the study of Martinez-Cedillo et al., the scene-viewing task was restricted to the retention interval of the memory task, meaning that monitoring for memory stimuli would not disrupt scene viewing. Following on from our previous research, the question naturally arises just how impervious to interference the social prioritisation effect is. In this context, it is important to realise that not all secondary tasks are equivalent.

According to load theory (e.g., Lavie et al., 2003, 2004), cognitive load may be contrasted with perceptual load. Cognitive load refers to a load upon the central mechanisms responsible for maintaining task priorities or task sets (see Moore et al., 2017), and may be impacted by a verbal working memory load. In contrast, perceptual load refers to the load placed on the resources required for perceptual analysis and identification of stimuli. While perceptual load is often manipulated by increasing the number or complexity of the stimuli in a display (see Lavie & Cox, 1997; Lavie & Tsal, 1994), Lavie and colleagues have suggested that it may also be increased by a secondary task in which the to-be-remembered stimuli come from the same domain as the primary stimuli (i.e., visuo-spatial memoranda in the case of viewing visual scenes). In particular, Konstantinou et al. (2012), Konstantinou and Lavie, 2013, Konstantinou et al. 2014 have shown that a visuo-spatial memory load can act to reduce perceptual processing in a way similar to directly increasing the number of stimuli. Typically, increased perceptual load results in reduced capture by irrelevant stimuli, but also more generally reduced perceptual processing of presented stimuli.

In the context of manipulations of perceptual load using display complexity or display size, it has been shown that certain stimuli like faces (e.g., Lavie et al., 2003) or objects of high familiarity (e.g., musical instruments for expert musicians; Ro et al., 2009) can continue to exert interference when irrelevant, even in the face of high perceptual load. Thus, it is certainly the case that the processing of irrelevant stimuli may continue to proceed even in the face of a high perceptual load. However, measuring the perceptual load of primary task visual stimuli is not straightforward (for discussion, see Roper et al., 2013; Tsal & Benoni, 2010). By some measures, complex and cluttered scenes may be considered to already present a high perceptual load. In order to side-step the issue of measuring the perceptual load imposed by scenes, in the present study we manipulated perceptual load by virtue of a secondary visuo-spatial memory load. Will the social prioritisation effect in scene viewing show invulnerability to an increased perceptual load imposed by a competing visuo-spatial memory task?

Other authors have recently argued that looking at people in scenes is automatic and reflexive (e.g. Rösler et al., 2017), since it occurs without an explicit goal to inspect people, it occurs even for the very first saccades, and it occurs in response to briefly presented images. Aside from the control of eye movements, earlier work also demonstrates that participants are able to detect the presence of animal and human bodies, and faces in complex images both with very brief displays (Kirchner, & Thorpe, 2006; Thorpe et al., 1996) and with attention diverted (Li et al., 2002). Such rapid identification may occur on the basis of the "feedforward" processing courtesy of specialised structures tuned to detect real objects (e.g., Van Rullen, 2009) and their constituent features (e.g., Evans & Treisman, 2005). If people in scenes are detected rapidly, then a key question is whether participants will fixate them regardless of the current task demands. This provides one test of automaticity, in the sense that if participants cannot modulate their prioritization of social regions then it is automatic. Of course, other aspects of processing may be important for automaticity (such as how rapidly they occur), and tasks need not be thought of as dichotomously fully automatic or completely under voluntary control. Our investigation contributes to this by testing whether orienting to people in scenes will continue undiminished, even in the face of competing demands from a second demanding task presented in the same sensory domain.

The present study examined the effect of increasing visuospatial memory load on image viewing. Is the social prioritisation effect sufficiently obligatory and automatic as to persist in the face of competition from this type of load? It is possible to think of the competition imposed by the working memory task as competition for shared perceptual processing resources. According to this view, increasing visuo-spatial working memory load would result in increased perceptual load, and would withdraw perceptual processing resources from the scene viewing task (see Konstantinou et al., 2012,

2014, for behavioural and neuroscientific evidence). To the extent that the explicitly relevant memory task outcompetes the scene-viewing task for any perceptual resources required to drive social prioritisation, this prioritisation should be diminished. In contrast, if either the social prioritisation effect does not depend on perceptual resources that overlap with the memory task, or the priority for processing social items is set so high as to outcompete the memory task, we may anticipate no effect.

An alternative way of conceptualising the competition that may occur between the two tasks is in terms of more direct competition to the distinct locations occupied by the stimuli in the two tasks. Retaining spatial information in memory may impact shifts of spatial attention directly, as eye movements to remembered locations may be recruited in the service of memory rehearsal (e.g., see Baddeley, 1986). Certainly, it has been known for some time that eye movements to locations that are incongruent with to-be-remembered locations result in interference with memory (e.g., Hale et al., 1996; Lawrence et al., 2004; Postle, 2006; Pearson & Sahraie, 2003). In addition, participants often make spontaneous eye movements to the locations of to-be-remembered stimuli, even for stimuli where the location is not explicitly relevant (e.g., Richardson & Spivey, 2000). However, it is not clear that these location-congruent eye movements are functional in the sense of improving memory. It is certainly the case that participants can maintain a high level of spatial memory even when fixation is held constant during a retention interval. For example, Smyth and Scholey (1994) showed that memory for a sequence of locations in the Corsi Blocks test was no different when participants were free to move their eyes in the retention interval, compared to when fixation was held constant (see also Godjin & Theeuwes, 2012, for a similar result). However, it is important to note that under these conditions shifts in spatial attention and eye-movement planning remain possible. The deleterious effect of irrelevant incongruent eye movements is then explained not by the eye movements per se, but by the shifts of spatial attention that typically accompany them. Thus, one possibility is that in the face of a high spatial working memory load participants will make fewer eye movements, and correspondingly fewer shifts of covert spatial attention to the scene during the retention interval in order to protect memory. Our question is then if the eye movements that underlie the social-prioritisation effect persist even though they may impose a cost to memory.

Research shows that if eye movements to memoranda during the Corsi Blocks task are not merely suppressed but rendered impossible by presenting the stimuli to participants with their head and eyes turned in opposite directions, such that the eyes cannot rotate any further to fixate the stimuli, there are costs to spatial memory performance (Ball et al., 2013; Pearson et al., 2014). However, here both actual eye movements, eye-movement planning, and exogenous deployments of attention (see Pearson et al., 2014) are all disrupted. These studies serve to underline the importance of possible eye movements and shifts of attention to relevant stimuli during the Corsi Blocks task.

Regardless of how one conceptualises the competition between the two tasks either mediated by perceptual load or more directly in terms of competition between the stimulus locations involved, the question remains whether the social prioritisation effect is sufficiently obligatory and automatic to remain undiminished in the face of such competition. The current study sought to answer this question.

A previous study by Cronin et al. (2020) compared the effects of a visual and a verbal memory load on scene viewing in general but did not specifically examine social prioritisation or the fixation of particular regions. The results showed that both a visual and a verbal memory load altered viewing patterns in the scene-viewing task, such that participants increasingly fixated on the centre of the screen. However, like the study of Bianchi et al. (2020), a no-load control was compared against a high-load condition, meaning it was impossible to separate the specific effect of increasing memory load from the more general effect of carrying out a second task at all. The current work went beyond that of Cronin by specifically comparing a small and large visuospatial memory load and specifically investigating the social prioritisation effect.

One important consideration regarding spatial working memory is the mode of presentation. Most early studies presented locations sequentially and asked participants to recall the items in the correct order so as to aid comparisons with verbal working memory, which is more naturally sequential. A number of studies have compared memory for simultaneously presented matrix patterns ("pattern span") with variations of the Corsi Blocks task in which participants attempt to recall sequentially presented locations, in the correct serial order. Della Salla et al. (1999), for example, showed that these two tasks showed distinct patterns of interference. Other research points to distinct developmental trajectories for these two tasks (e.g., Logie & Pearson, 1997; Pickering et al., 2001). Critically, recent research using the abducted eye paradigm, in which saccades and exogenous shifts of spatial attention are rendered impossible, showed that sequential spatial memory tasks were interfered with but simultaneous spatial memory tasks were not. These data thus suggest that sequential tasks overlap to a greater extent with mechanisms for shifts of attention and eye movements. Recent research (e.g., Cortis Mack et al., 2015; Gmeindl et al., 2011) shows that even when serial order is not required participants tend to recall spatial locations in the order in which they were presented, at least for short lists of locations. Thus, whether or not participants are required to retain sequential spatial locations in order or not may make little

difference to the underlying mechanisms involved at least for short lists of locations. On balance, previous research suggests that sequentially presented sets of spatial locations likely recruit shifts of spatial attention as a rehearsal mechanism to a greater extent than simultaneously presented sets of locations. Simultaneously presented sets of locations, in contrast, likely recruit configural representations that capture the overall form of the set of locations presented (e.g., see Gmeindl et al., 2011; Jiang et al., 2000).

Given the potentially greater role for spatial attention in maintaining sequentially presented spatial locations in memory, we compared simultaneously and sequentially presented spatial locations. If any disruption of the social prioritisation effect can be tied to competition for spatial attention and associated fixations, rather than more general perceptual resources, then disruption should be greatest in the sequential presentation condition.

Method

Participants

Sixty participants (ages 18–32 years, *M = 21.6 SD = 4.85* years; 42 females, 18 males) from the University of Essex took part in the experiment. Half of these participants took part in the sequential condition and half in the simultaneous condition, which were run at separate times. After discarding data from 13 participants who were not accurate in the calibration (mean error above 0.8°, a threshold set a priori) the sample size was 47. All participants reported normal or corrected-to-normal vision. Participants were paid £4 or received course credit for their participation.

Apparatus and stimuli

The experiment was programmed in MATLAB (version 9.1.0, R2016b; the Mathworks, Natick, MA, USA), using the Psychophysics Toolbox. Eye position was recorded using the SMI RED500, which is a screen-based eye tracker that samples pupil position at 500 Hz. A 9-point calibration and validation were repeated several times to ensure that all recordings had a mean spatial error of better than 0.8° (mean across participants = 0.53°, SD = 0.27). Head movements were restricted using a chin rest. The experiment took place in a dimly illuminated, sound-attenuated room. Participants sat 60 cm away from the screen (a 22-in. CRT monitor) so that the stimuli subtended approximately 43° by 28° of visual angle at 1,680 × 1,050 pixels. The stimuli were a set of 64 high-resolution colour photographs, which were found from free access image databases, for instance Pixabay. Thirty-two of these were filler photographs that we did not analyse further but were used to divert attention from the

experimental images. The fillers were taken from a similar set of images but did not contain the key objects of interest. The remaining 32 experimental images were selected to contain a pre-specified area of interest on opposite sides of the image: a social area (a person) and a non-social area (an object). We created two versions of each image using Pic-Monkey, an online image editing software, by manipulating the non-social object so that it was either high or low in visual saliency. These manipulations made the object stand out from its background to a greater or lesser degree, for example by changing the colour, and this was confirmed by using an implementation of the Itti and Koch saliency map model (Itti & Koch, 2000) via MATLAB (version 9.1.0, R2016b). This model selected the high-saliency object early (the first three simulated fixations) and the low-saliency object on later simulated fixations. Further details of these stimuli can be found in Martinez-Cedillo et al. (2022).

Procedure

Figure 1 summarises the experimental procedure. Participants performed a visuospatial working memory task where they were required to remember a set of dots. While doing this, they viewed each of the scenes.

The calibration and validation of the eye tracker were performed at the start of the experiment. The working memory task was adapted from McNab et al. (2008). Each trial started with a fixation dot displayed for 500 ms. After this, a 4 × 4 grid was displayed with either one (low load) or six (high load) dots presented in different locations on the grid. The dots were green on a black background.

In the sequential presentation condition, dots appeared one at a time (in high load). Each location was presented for 333 ms before disappearing and being replaced by the next location. In the low-load condition, the single location was presented for 2,000 ms. These timings were designed to ensure that the total presentation time was the same in both high- and low-load conditions (see McNab et al., 2008). In the simultaneous presentation condition, all the dots were shown on the screen at the same time for 2,000 ms. In all cases, the locations were chosen at random with no location repetition and followed by a fixation point presented for 500 ms.

Next, the scene was shown for 5,000 ms. Participants were instructed to look freely at the scene while performing the memory task. After each scene, a probe display was presented, and participants were required to indicate whether the probe location was one of those presented previously (or not) by making a keyboard response. After each response, a feedback screen was displayed, showing the reaction time and accuracy from that trial. Participants were encouraged to respond as quickly and accurately as possible.
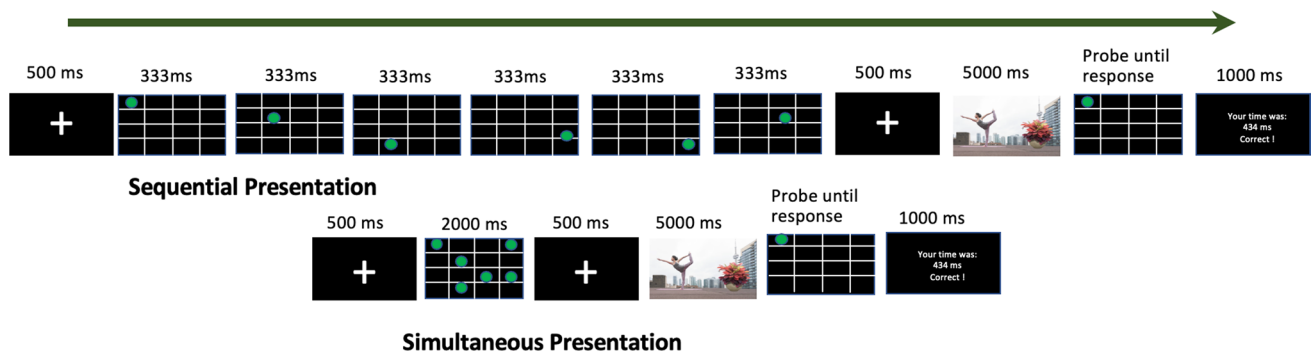
Fig. 1 Experimental procedure showing the sequential (**above**) and simultaneous (**below**) memory task. Both examples show a trial from the high-load condition

The experiment consisted of two blocks with low (one dot) or high (six dots) memory load. Each block consisted of 32 trials (half fillers and half experimental images). Half of the participants started with the low-load block and half with the high-load block. Trials were equally divided between those with high and low saliency objects, and between those where the memory probe was present or absent. Experimental images were counterbalanced across participants such that each scene appeared in all load and saliency conditions, and scenes were also mirror reversed for half the participants to control for any biases to the left or right of the image. The experiment took approximately 25 min.

Data analysis

Participants who scored below 50% on the memory task were excluded from the analysis. Trials with incorrect memory responses, and those in which the fixation at the onset of the scene was not on the centre, were also removed. Fixations were removed if their duration was below 100 ms. Following these criteria, we analysed data from 44 participants (ages 18–32 years, $M = 21.43, SD = 4.78$ years; 34 females, ten males; 20 in sequential and 24 in simultaneous). For our main analyses, we compared participant means using a mixed ANOVA with the factors of load (high and low) and object saliency (high and low) and a between-subjects' factor of presentation type (sequential and simultaneous). A sensitivity analysis indicated that our sample size allows us to detect medium effect sizes (Cohen's $f = 0.65$) with a power of at least 95% when assuming a correlation of $r = .50$ between factor levels.

Results

We examine the effects of encoding visuospatial information on the image-viewing task. We first report behavioural data from the memory task and general eye-movement behaviour.

Then, we examine the effects of load on fixations to social and non-social elements.

Behavioural data

Accuracy in the memory task was lower in the high-load condition *(M = 78.56%, SD = 21.47)* than in the low-load condition *(M = 90.93%, SD = 11.09)*. A paired-sample t-test compared the accuracy of the memory probe under high and low loads, *t* (43) = -3.419, *p* < .001, confirming that the high-load task was more difficult.

General eye-movement statistics

To investigate whether load interfered with eye movements, we examined the number and duration of fixations and the degree to which they showed a bias to the centre of the screen. This was compared across load and saliency conditions, but also across the two different types of presentation since we reasoned that controlling the sequence of the dots (sequential condition) might have an impact on eye movements during memory retention. Table 1 summarises these general eye-movement statistics.

Although there were some small differences between the number of fixations per trial and the average fixation duration, there were no significant effects of saliency, load or presentation type on either of these measures *(all F<1)*. The mean distance between each fixation and the centre of the image quantifies the extent to which participants were biased to look in the centre, with lower average distances indicating a greater central bias. If the memory task affects this bias (e.g., because participants adopt a strategy of staring at the centre), then we might expect a difference in this measure between low and high loads. We did indeed find a main effect of load on participants' distance to the centre, $F (1, 42) = 4.759, p = 0.035, \eta^2 = 0.102$. There were no significant effects of saliency or presentation type (both

Table 1 Measures of general viewing behaviour across conditions. Values show the mean across participants, with standard deviation in parentheses

	Load	High load		Low load	
	Saliency of non-social object	High	Low	High	Low
	Presentation type				
Number of fixations per trial	Sequential	18.49 (6.12)	16.75 (6.52)	19.06 (5.24)	18.54 (6.63)
	Simultaneous	17.15 (7.88)	16.67 (5.61)	17.03 (6.58)	17.64 (5.04)
Average fixation duration (ms)	Sequential	164.76 (64.00)	239.32 (117.42)	172.66 (43.40)	232.91 (100.59)
	Simultaneous	165.97 (67.36)	162.66 (62.70)	164.76 (64.00)	172.66 (43.40)
Average distance to the centre (degrees)	Sequential	11.19 (1.7)	11.31 (1.91)	11.29 (2.89)	11.97 (3.35)
	Simultaneous	10.55 (3.25)	10.12 (2.7)	11.46 (2.45)	11.89 (1.85)
Heatmap entropy (bits)	Sequential	18.91 (0.56)	18.82 (0.68)	19.24 (0.79)	19.43 (0.64)
	Simultaneous	18.68 (0.79)	18.72 (0.67)	19.45 (0.83)	19.54 (0.70)

$F<1$), and no load and presentation type interaction, $F (1, 42) = 1.463$, $p = 0.233$, $\eta^2 = 0.034$, or higher order interactions. This indicates that there is an increased centre bias in the high-load condition relative to the low-load condition.

The same difference is visible when looking at the overall fixation distribution (Fig. 2). These distributions were created by convolving the location of all fixations in each condition with a Gaussian kernel. We also generated fixation maps for each participant, and we used the heatmap entropy to quantify how spread out or exploratory the fixations were in each condition. Entropy is a measure from information theory that has often been used to quantify the dispersion of eye fixations (for a review, see Shiferaw et al., 2019). Entropy varies from a theoretical minimum of zero (if all fixations were on exactly the same pixel) to a maximum that depends on the size of the image. Distributions with higher entropy are more dispersed or less predictable. In Table 1 we report the mean (and SD) entropy in bits. There was a significant effect of load on entropy, $F (1, 42) = 42.272$, $p < 0.001$, $\eta^2 = 0.502$. Fixations in low-load trials were more widely dispersed across the viewing area than those in the high-load condition. There were no effects of saliency or presentation type, and no interactions (all $F<1$).

Fixations to social elements

Table 2 shows the proportion of fixations to the different regions of interest. It is clear from this table that, across all conditions, the social area was fixated much more often than the non-social area. We therefore examined the proportion of fixations on the social area in order to see whether memorising

spatial information affects this social prioritisation effect. There was a significant effect of load, $F (1, 42) = 5.156$, $p = 0.029$, $\eta^2 = 0.114$, indicating that participants looked less often at the social object when memorising high loads of
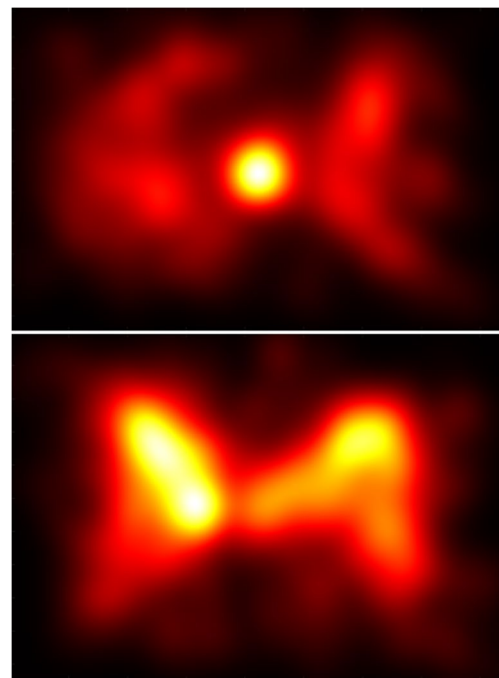

Fig. 2 Fixation distribution across participants (pooled over sequential and simultaneous presentation) for high load (**above**) and low load (**below**). Hotter colours indicate a greater density of fixations. Note that fixations were less spread out during the high-load condition (entropy = 18.78) than during the low-load condition (entropy = 19.42)

Table 2 Percentage of fixations on each region of interest. Values show the mean across participants, with standard deviation in parentheses

	Load	High load		Low load	
	Saliency of non-social object	High	Low	High	Low
	Presentation type				
Percentage fixations on the social area	Sequential	39.37% (21.71%)	32.25% (21.47%)	44.34% (14.85%)	39.53% (15.68%)
	Simultaneous	34.51% (20.39%)	37.80% (23.37%)	43.14% (10.37%)	45.46% (11.62%)
Percentage fixations on the non-social area	Sequential	24.29% (15.01%)	18.11% (13.06%)	24.57% (18.11%)	17.24% (10.90%)
	Simultaneous	20.13% (13.96%)	18.73% (13.08%)	23.32% (12.67%)	23.96% (11.22%)

information (Fig. 3). There was no interaction between load and presentation type, $F(1, 42) = 0.195$, $p = 0.661$, $\eta^2 = 0.005$. There was no main effect of saliency or presentation type (both $F<1$), but a marginal interaction between saliency and presentation type, $F(1, 42) = 3.623$, $p = 0.064$, $\eta^2 = 0.079$. No other interactions were significant (all $F<1$).

Fixations to non-social elements

For this analysis, we considered the proportion of fixations on the non-social. In particular, we reasoned that higher loads might make people more susceptible to capture by salient objects, resulting in a larger difference between high and low salient objects in the high-load condition. There was a trend towards an effect of saliency, $F(1, 42) = 3.695$, $p = 0.061$, $\eta^2 = 0.081$, with high salient objects fixated more often. However, there was no effect of load ($F<1$), presentation type ($F<1$), or their interaction ($F<1$). The effect of saliency was more noticeable in the sequential presentation condition, but the interaction between saliency and presentation type fell short of significance, $F(1, 42) = 2.890$, $p = 0.097$, $\eta^2 = 0.064$.

The first fixation

In previous studies (e.g., Anderson & Donk, 2017) it has been proposed that the first fixation on a scene might reflect "early", more automatic attentional capture. We therefore carried out an additional analysis on the landing point of the first saccade (i.e., the first fixation after leaving the centre where viewing was constrained to begin). We examined the proportion of these fixations on the social and non-social areas in each condition. For the proportion of fixations to the social object, there was an effect of load that matched the pattern seen in the overall analysis, $F(1, 42) = 6.60$, $p = 0.014$, $\eta^2 = 0.136$. There were fewer initial fixations to the social area in the high-load condition ($M = 36.26\%$, $SEM = 4.21$) than in the low-load condition ($M = 49.25\%$, $SEM = 3.56$). In general, it was quite likely that the first saccade was targeted towards the social region of interest. There were no significant effects of saliency ($F<1$) or presentation type ($F<1$), and no interactions. For the proportion of first
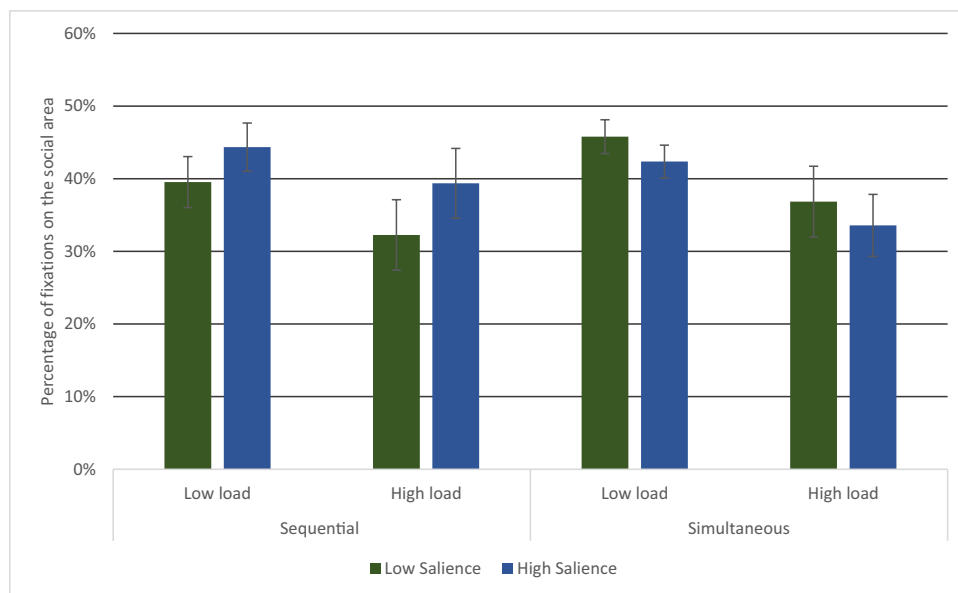

Fig. 3 Percentage of fixations on the social area, across conditions. Bars show the mean with standard errors

fixations to the non-social object, there was a marginal effect of saliency, $F(1, 42) = 3.11$, $p = 0.085$, $\eta^2 = 0.069$. High saliency objects ($M = 13.11\%$, $SEM = 1.81$) were somewhat more likely to be fixated than low saliency objects ($M = 8.71\%$, $SEM = 1.77$), but these fixations were quite rare in general. There were no other significant effects on the first fixations to the non-social object (all $F<1$).

General discussion

We asked participants to memorise spatial information while freely viewing scenes containing social and non-social items. Our aim was to examine whether the social prioritisation effect would remain undiminished in the face of competition for perceptual processing resources, or competing demands on eye movements and attention, in the two tasks. In general, observers spent much of their time fixating on the social object in our pictures, consistent with previous reports (Birmingham et al., 2009; End & Gamer, 2019; Martinez-Cedillo et al., 2022). In all conditions, the social object was fixated much more often than the non-social object. This was also the case on the very first fixation, with up to 50% of the first saccades targeting the person in the scene. In addition, we found little effect of the salience of the non-social object, suggesting that in the face of the presence of the social item, the salience of the non-social elements no longer exerts a strong effect on performance (see Birmingham et al., 2009).

In previous research (Martinez-Cedillo et al., 2022) we showed that the social prioritisation effect was immune to competition from increasing the demand imposed by a secondary cognitive task. This invulnerability to disruption by "cognitive load" contrasts with the increased interference typically seen when cognitive load is increased in the context of physically salient distractors (Boot et al., 2005; Burnham, 2010; Lavie & De Fockert, 2005). These prior results suggested that the social prioritisation effect may be automatic, in the sense that the implicit goal to attend to social elements overrides or is distinct from the control processes used to retain information in working memory.

Critically, in the present study, fixations to the social object *were* reduced when participants retained six compared to one location in memory. This is a novel effect. Although the social region was still fixated more than the non-social region, even in high-load trials, it indicates that, at least some of the time, participants were able to avoid looking at people in conditions of high load. The same effect of load was found on the first eye movement in the scene, with this eye movement being more likely to target the social object when memory load was low. The overall number of eye movements did not differ as a function of load. Instead, the distribution of eye movements between the social element and the other areas changed. One reason for this is that

participants tended to look more towards the centre of the screen, and to be less exploratory, in the high-load condition (see below for a further discussion of this point).

Fixations to the non-social object were not affected by load. This is interesting because it suggests that high-load conditions did not only cause participants to look at the centre of the screen and avoid looking at anything; instead, the load effect was selectively on the social region, which implies that orienting to people requires different resources when it comes to attentional control. The effects of model-predicted saliency, meanwhile, were limited, with only a marginal difference between high and low saliency objects, even on the first saccade. The present study was not constructed to look at this in detail, and it may be that the changes in saliency here were not large enough to have an impact, or that they were masked by other confounds in the choice of stimuli. In future research it would be possible to look at the predictiveness of saliency in different load conditions, across the whole image, and with reference to more recent modelling approaches (e.g., Kümmerer et al., 2015).

The current results place some limits on the social prioritisation effect, showing certain conditions under which it is reduced. Other researchers have shown that social prioritisation can be modified by instructions that require inspecting different parts of a scene. For example, Flechsenhar and Gamer (2017) asked participants to perform different tasks such as counting blue objects or estimating the percentage of white pixels in a scene. As in the present study, people in the scenes continued to be fixated more than non-social and salient objects, even though the people were not relevant for the task. As we would expect, the bias to look at people could be reduced by a task explicitly focusing on other parts of the image. However, the role of the difficulty of the different tasks was not clear in that study. In the present experiment, in contrast, we used a secondary memory task that did not involve the inspected images and in which we could control task difficulty through load.

In the *Introduction*, we described two ways of thinking about the potential impact of a spatial memory load on the social prioritisation effect. According to load theory, the identification of visual stimuli draws on perceptual resources that are shared between competing visual stimuli and tasks. In particular, the theory suggests overlap between the resources required to complete visual and spatial working memory and perceptual tasks. It may be that under such conditions the social elements are not fully identified and therefore not tagged as social elements of particular relevance. An alternative is that the social elements are identified as social but under conditions of competition they do not draw spatial attention to their location, either due to resource limitations or because such spatial attention is being used to rehearse locations in spatial working memory (see Awh & Jonides, 2001; Smyth & Scholey, 1994).

Related to this latter possibility it could have been that those eye movements to areas identified as socially relevant are suppressed and overridden in order to prevent interference with shifts of spatial attention that serve to rehearse locations in the spatial memory task. However, arguing against this possibility we found equivalent disruption of the social prioritisation effect by both simultaneous and sequential presentation of the spatial memory stimuli. Previous research has drawn a distinction between simultaneous and sequential presentation; for example, the model of Logie and Pearson (1997) distinguishes the "visual cache", which is concerned with the configural properties of sets of spatial locations, with the "inner scribe", which acts to serially rehearse sets of sequentially presented locations. Indeed, empirical data from dual-task studies (e.g., Della Salla et al., 1999) and from performance in the abducted eye paradigm (Pearson et al., 2014) supports the greater involvement of shifts of the eyes and attention in supporting sequential rather than simultaneous spatial memory. That we observe equivalence between simultaneous and sequential spatial memory argues against the idea that the reduction in the social prioritisation effect observed is linked to serial spatial rehearsal.

The most parsimonious explanation of the current data is in terms of more general competition for perceptual resources leading to reduced processing of the scene. This competition had selective effects on the social objects. Pushing this explanation to its logical conclusion, it might suggest that the identification of significant social objects in scenes is not necessarily resource-free. Earlier suggestions that processing faces and highly familiar objects might be immune to increases in task-relevant perceptual load could then be traced to the use of relatively simple displays that are not characteristic of typical human environments. The picture that then emerges of the social prioritisation effect in scene viewing suggests that the effect depends on the availability of perceptual processing resources but not more general cognitive resources related to working memory. However, exactly how to explain the social prioritisation effect remains to be determined. Given that social and non-social objects do not seem to be differentially easy to categorise in scenes (e.g., Li et al., 2021, showed that both animals and vehicles could be rapidly categorised in scenes), it may be that the representations of social objects are instead tagged in memory as having high levels of importance, such that once identified they go on to have priority in controlling shifts of attention.

Declarations

References

Anderson, N. C., & Donk, M. (2017). Salient object changes influence overt attentional prioritization and object-based targeting in natural scenes. *PLoS One, 12*(2), e0172132.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in cognitive sciences, 5*(3), 119–126.

Baddeley, A. D. (1986). *Working memory oxford*. Oxford University.

Ball, K., Pearson, D. G., & Smith, D. T. (2013). Oculomotor involvement in spatial working memory is task-specific. *Cognition, 129*(2), 439–446.

Bianchi, L. J., Kingstone, A., & Risko, E. F. (2020). The role of cognitive load in modulating social looking: a mobile eye tracking study. *Cognitive research: principles and implications, 5*(1), 1–21.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Gaze selection in complex social scenes. *Visual Cognition, 16*, 341–355.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Social attention and real-world scenes: The roles of action, competition and social content. *Quarterly Journal of Experimental Psychology, 61*, 986–998.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision research, 49*(24), 2992–3000.

Boot, W. R., Kramer, A. F., & Peterson, M. S. (2005). Oculomotor consequences of abrupt object onsets and offsets: Onsets dominate oculomotor capture. *Perception & Psychophysics, 67*(5), 910–928.

Burnham, B. R. (2010). Cognitive load modulates attentional capture by color singletons during effortful visual search. *Acta psychologica, 135*(1), 50–58.

Cortis-Mack, C., Dent, K., Kennett, S., & Ward, G. (2015). First things first: similar list length and output order effects for verbal and

nonverbal stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(4), 1179.

Cronin, D. A., Peacock, C. E., & Henderson, J. M. (2020). Visual and verbal working memory loads interfere with scene-viewing. *Attention, Perception, & Psychophysics, 82*(6), 2814–2820.

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: A tool for unwelding visuo–spatial memory. *Neuropsychologia, 37*(10), 1189–1199.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193–222.

Donk, M., & Van Zoest, W. (2008). Effects of salience are short-lived. *Psychological Science, 19*(7), 733–739.

End, A., & Gamer, M. (2017). Preferential processing of social features and their interplay with physical saliency in complex naturalistic scenes. *Frontiers in Psychology, 8*, 418.

End, A., & Gamer, M. (2019). Task instructions can accelerate the early preference for social features in naturalistic scenes. *Royal Society open science, 6*(3), 180596.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1476.

Flechsenhar, A. F., & Gamer, M. (2017). Top-down influence on gaze patterns in the presence of social features. *PLoS One, 12*(8), e0183799.

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of vision, 8*(2), 6–6.

Gmeindl, L., Walsh, M., & Courtney, S. M. (2011). Binding serial order to representations in working memory: A spatial/verbal dissociation. *Memory & cognition, 39*(1), 37–46.

Godijn, R., & Theeuwes, J. (2012). Overt is no better than covert when rehearsing visuo-spatial information in working memory. *Memory & Cognition, 40*(1), 52–61.

Hale, S., Myerson, J., Rhee, S. H., Weiss, C. S., & Abrams, R. A. (1996). Selective interference with the maintenance of location information in working memory. *Neuropsychology, 10*(2), 228.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision research, 40*(10–12), 1489–1506.

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

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision research, 46*(11), 1762–1776.

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.

Konstantinou, N., Bahrami, B., Rees, G., & Lavie, N. (2012). Visual short-term memory load reduces retinotopic cortex response to contrast. *Journal of Cognitive Neuroscience, 24*(11), 2199–2210.

Konstantinou, N., Beal, E., King, J. R., & Lavie, N. (2014). Working memory load and distraction: dissociable effects of visual maintenance and cognitive control. *Attention, Perception, & Psychophysics, 76*(7), 1985–1997.

Kümmerer, M., Wallis, T. S., & Bethge, M. (2015). Information-theoretic model comparison unifies saliency metrics. *Proceedings of the National Academy of Sciences, 112*(52), 16054–16059.

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., & De Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic bulletin & review, 12*(4), 669–674.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & psychophysics, 56*(2), 183–197.

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.

Lavie, N., Ro, T., & Russell, C. (2003). The role of perceptual load in processing distractor faces. *Psychological science, 14*(5), 510–515.

Lawrence, B. M., Myerson, J., & Abrams, R. A. (2004). Interference with spatial working memory: An eye movement is more than a shift of attention. *Psychonomic Bulletin & Review, 11*(3), 488–494.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Detection of objects in natural scenes with minimal or no attention. *Journal of Vision, 2*(7), 250–250.

Li, W., Guan, J., & Shi, W. (2021). Increasing the load on executive working memory reduces the search performance in the natural scenes: Evidence from eye movements. *Current Psychology*, 1–14.

Logie, R., & Pearson, D. G. (1997). The inner eye and the inner scribe of visuo-spatial working memory: Evidence from developmental fractionation. *European Journal of cognitive psychology, 9*(3), 241–257.

Martinez-Cedillo, A. P., Dent, K., & Foulsham, T. (2022). Do cognitive load and ADHD traits affect the tendency to prioritise social information in scenes? *Quarterly Journal of Experimental Psychology, 75*(10), 1904–1918.

McNab, F., Leroux, G., Strand, F., Thorell, L., Bergman, S., & Klingberg, T. (2008). Common and unique components of inhibition and working memory: an fMRI, within-subjects investigation. *Neuropsychologia, 46*(11), 2668–2682.

Moore, D. J., Eccleston, C., & Keogh, E. (2017). Cognitive load selectively influences the interruptive effect of pain on attention. *Pain, 158*(10), 2035–2041.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly journal of experimental psychology, 11*(1), 56–60.

Pearson, D. G., Ball, K., & Smith, D. T. (2014). Oculomotor preparation as a rehearsal mechanism in spatial working memory. *Cognition, 132*(3), 416–428.

Pearson, D., & Sahraie, A. (2003). Oculomotor control and the maintenance of spatially and temporally distributed events in visuo-spatial working memory. *The Quarterly Journal of Experimental Psychology Section A, 56*(7), 1089–1111.

Pickering, S. J., Gathercole, S. E., Hall, M., & Lloyd, S. A. (2001). Development of memory for pattern and path: Further evidence for the fractionation of visuo-spatial memory. *The Quarterly Journal of Experimental Psychology Section A, 54*(2), 397–420.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139*(1), 23–38.

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: Looking at things that aren't there anymore. *Cognition, 76*(3), 269–295.

Ro, T., Friggel, A., & Lavie, N. (2009). Musical expertise modulates the effects of visual perceptual load. *Attention, Perception, & Psychophysics, 71*(4), 671–674.

Roper, Z. J., Cosman, J. D., & Vecera, S. P. (2013). Perceptual load corresponds with factors known to influence visual search. *Journal of experimental psychology: human perception and performance, 39*(5), 1340.

Rösler, L., End, A., & Gamer, M. (2017). Orienting towards social features in naturalistic scenes is reflexive. *PLoS One, 12*(7), e0182037.

Shiferaw, B., Downey, L., & Crewther, D. (2019). A review of gaze entropy as a measure of visual scanning efficiency. *Neuroscience & Biobehavioral Reviews, 96*, 353–366.

Smyth, M. M., & Scholey, K. A. (1994). Interference in immediate spatial memory. *Memory & Cognition, 22*(1), 1–13.

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 22*(4), 1005.

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta psychologica, 135*(2), 77–99.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520–522.

Tsal, Y., & Benoni, H. (2010). Diluting the burden of load: perceptual load effects are simply dilution effects. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1645.

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology, 59*(11), 1931–1949.

Van Rullen, R. (2009). Binding hardwired versus on-demand feature conjunctions. *Visual cognition, 17*(1–2), 103–119.