Attention Scaling Modulates the Effective Capacity of Visual Sensory Memory

Leon Gmeindl†^{*a}, Lisa N. Jefferies^{*b}, and Steven Yantis^{a,c}

This is a post-peer-review, pre-copyedit version of an article published in Psychological Research. The final authenticated version is available online at: <https://doi.org/10.1007/s00426-018-1114-4>

Attention Scaling Modulates the Effective Capacity of Visual Sensory Memory

Leon Gmeindl^{†*a}, Lisa N. Jefferies^{*b}, and Steven Yantis^{a,c}

[†] Corresponding author; Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218-2686, USA; telephone: (410) 516-3357; electronic mail: gmeindl@jhu.edu

* Contributed equally

^a Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218-2686, USA

^b School of Applied Psychology, Griffith University, Southport, QLD, 4222, Australia

^c Posthumous

Acknowledgments

Funding was provided by a National Institute on Aging postdoctoral National Research Service Award (T32-AG027668) to L.G., a New Researcher Grant from the Menzies Health Institute Queensland to LNJ, and a National Institutes of Health grant R01-DA013165 to S.Y. The authors thank Howard Egeth for providing comments on a previous draft of this paper, and Michelle Hall and Taylor Chamberlain for research assistance.

Running Head: ATTENTION AND SENSORY MEMORY CAPACITY

Visual sensory memory (VSM) has a high capacity, but its contents are fleeting. Recent evidence that the breadth of attention strongly influences the efficiency of visual processing suggests that it might also modulate the effective capacity of VSM. We manipulated the breadth of attention with different cue sizes and used the partial-report technique to estimate the capacity of VSM. Whether attention was deployed voluntarily or captured by a salient cue, narrowly focused attention increased the effective capacity of VSM. This study reveals, for the first time, a direct influence of the breadth of attention on the effective capacity of VSM.

Keywords: sensory memory; iconic memory; attention capture; voluntary attention; attention focus

1. Introduction

More than half a century ago, Sperling (1960) demonstrated that people encode and temporarily store much more information from a brief visual presentation than they can report only a few hundred milliseconds later. Visual sensory memory (VSM, also called “iconic memory;” Neisser, 1967) thus has a high capacity but its contents are short-lived. Despite or perhaps because of its ephemeral nature, VSM provides an abundant, frequently-updated supply of recent stimulus information from which more durable representations can be established, maintained, and manipulated in working memory. As perhaps the earliest stage of visual memory, VSM occupies an important position at the interface between the external world on the one hand and working memory and cognition on the other. As such, it is important to understand how much information can be held in VSM and what factors may influence its capacity. The present study specifically examined whether changes to the spatial extent of selective attention modulate the effective capacity of VSM.

Sperling’s (1960) partial-report technique (see also Averbach & Coriell, 1961) revealed that most of the stimulus information stored in VSM can be accessed shortly after removal of the stimuli. In Sperling’s experiments, participants were given brief (e.g., 50-ms) presentations of alphanumeric characters arrayed in a grid formation (e.g., 3 rows, 4 columns). In one condition, a partial-report tone was presented after the characters were removed, and its pitch (e.g., high, middle, or low) indicated which row of characters participants should report. With this sampling technique, the total number of characters held in VSM could be estimated based on the proportion of items correctly reported from each cued row. Sperling found that whereas participants could accurately

recall only about four characters from the entire array when no partial-report tone was provided (perhaps consistent with the capacity of working memory; see Cowan, 2001), the estimated capacity of VSM (i.e., the number of characters stored in memory that were available for retrieval) was at least twice as large when a partial-report tone was provided. This tone directs attention selectively to a subset of stimulus representations available in VSM, which is then encoded in working memory and therefore available for report. Together with a substantial body of knowledge gained from psychophysical studies of the properties and substages of VSM (see Coltheart, 1980; Irwin & Thomas, 2008; and Long, 1980, for reviews), these findings motivate further investigation of the interactions between selective attention and VSM.

Of course, selective attention might play an important role not only in the retrieval of information from VSM evoked by the partial-report cue, but also (and perhaps even more fundamentally) in the quality and duration of VSM representations, which would translate into a change in the effective capacity of VSM. Previous studies have shown that narrowly-focused visuospatial attention (as compared to broadly-distributed attention), when allocated to the region of space encompassed by a small rather than a large spatial pre-cue, improves the detection and discrimination of stimuli appearing within the cued region (e.g., Castiello & Umiltà, 1990; Eriksen & St. James, 1986; Eriksen & Yeh, 1985; Maringelli & Umiltà, 1998; Turatto et al., 2000) as well as influencing at the earliest stages of attention-modulated brain activity (e.g., Fu et al., 2005; Luo, Greenwood, & Parasuraman et al., 2001; Wang et al., 2010), even when the location of the stimulus is known in advance. This *cue-size effect* has been replicated many times and indicates that the efficiency of target detection and identification is

inversely related to the breadth of attention. These data led us to test whether the spatial extent of the attentional focus (narrow or broad) at the time of stimulus presentation affects the quality and/or duration of VSM representations. This, in turn, may affect how much information can be maintained in VSM – in other words, the breadth of attention may modulate the effective capacity of VSM.

Experiment 1 consisted of two parts which test whether reflexive changes to the breadth of attention modulate the effective capacity of VSM. In Part 1, small or large abrupt-onset cues were used to capture attention to either a relatively narrow or broad region of space, respectively. In order to combine the use of cues with Sperling's partial-report technique, in which the tone indicates which stimulus row to report, both the small and large cues encompassed the entire stimulus array. In Part 2, we test whether preventing the automatic rescaling of attention to the cues would eliminate the effect of the cues on VSM. Thus, Part 2 provides an important control condition to confirm that any changes to the effective capacity of VSM observed in Part 1 stem from changes to the breadth of attention, not from other effects such as alerting. We attempted to prevent the automatic rescaling of attention elicited by abrupt-onset pre-cues by displaying placeholder stimuli that explicitly marked the locations of impending task-relevant stimuli. This hypothesis builds upon the claim by Turatto et al. (2000) that the focus of attention can be adjusted (or maintained) voluntarily in the presence of a perceptual object to which attention can be directed, suggesting that the presence of placeholders that specifically mark the locations of the upcoming stimuli should prevent the automatic rescaling of attention. We refer to parts 1 and 2 of Experiment 1 as the *placeholders absent* and *placeholders present* conditions, respectively.

2. Experiment 1

2.1. Participants

Fourteen participants (M (SE) age = 22.7 (0.9) years; 8 females, 6 males) reporting normal hearing and normal or corrected-to-normal vision were included in the analyses reported below. An additional participant was excluded due to failure to maintain eye fixation on 12% of trials; the remaining participants failed to maintain fixation on fewer than 2% of trials. Participants were paid \$10/hour and were naïve as to the purpose of the experiment.

a

b

c

Fig. 1 *Panel a:* Schematic illustration of the placeholders absent condition in Experiment 1. After a randomly selected foreperiod of 500 or 1000 ms, either a small or large rectangular pre-cue appeared abruptly. After 150 ms, a 3 x 4 array of alphanumeric characters was presented for 50 ms. Immediately after the offset of the character array, a high, middle, or low partial-report tone indicated which row of the characters participants should recall. Stimuli are not drawn to scale. *Panel b:* Mean estimated capacity (number of characters available in VSM) as a function of pre-cue size in the placeholders absent condition. *Panel c:* Mean estimated capacity as a function of pre-cue size in the placeholders present condition. Error bars represent the standard error of the cue-size effect (i.e., difference in estimated capacity for large vs. small pre-cues)

2.2. Stimuli and apparatus

Stimulus presentation was controlled by custom Matlab (The MathWorks, Inc.) code using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). All visual stimuli were presented in black against a white background on a 17" CRT monitor (refresh rate = 120 Hz) in a dimly lit room. A chinrest ensured a viewing distance of 58 cm. A small filled rectangle ($\sim 0.06^\circ$ per side) remained at the center of the screen throughout the experiment and served as a fixation point (Fig. 1a). The small pre-cue was a black, unfilled rectangle 4.47° in height and 4.72° in width. The large pre-cue was a black, unfilled rectangle 9.13° in height and 9.38° in width. Both pre-cues had a border thickness of 0.26° and were centered on the fixation point. Twelve alphanumeric characters (each 0.70° in height and 0.51° in width, uppercase Monaco font) were presented in a 3-row $\times$ 4-column array (inter-character spacing $\sim 0.34^\circ$ horizontally and 0.26° vertically) centered on the fixation point; the entire array subtended $2.62^\circ \times 3.07^\circ$ and therefore was presented within the common region encompassed by both pre-cues (the edge of the character array was located $\sim 0.57^\circ$ horizontally and $\sim 0.67^\circ$ vertically from the border of the small pre-cue). An ensemble of characters was used, including digits 1 through 9 and the letters of the English alphabet, except for vowels and the letter Y (replicating Sperling, 1960, in order to minimize the chance of participants interpreting character arrays as words), as well as Q, S, and Z (to minimize confusion with the digits 0, 5, and 2), and the letter W (the only letter with a multisyllabic name). Within each row, two letters and two digits were presented, in randomized positions. No character could appear at more than one location within the array. The partial-report cue was a 50-ms tone approximately 2500 Hz (high), 750 Hz (middle), or 200 Hz (low) in pitch, played

through a computer loudspeaker. In order to ensure that participants maintained central fixation rather than making a saccade to the outline of the cue, the experimenter closely monitored the participant's eye position via a closed-circuit monitor connected to a high-contrast video camera focused tightly on the participant's right eye. If the experimenter judged that a saccade was made between the start of the trial and the sounding of the partial-report tone, an error message was displayed to the participant and the trial was excluded. In a previous study (Jefferies, Gmeindl, & Yantis, 2014), we directly compared the efficacy of detecting saccades both during and following the presentation of large and small cues with the same apparatus used in the present study (same experimenter, camera, and closed-circuit monitor) and with an EyeLink-1000 eye-tracker. The results showed the two methods to be highly comparable, with 1.9% and 2.4% of trials being excluded due to saccades, respectively ($t(29) < 1$, n.s.).

2.3. Procedure

In the placeholders absent condition, participants were instructed to fixate on the central filled rectangle (fixation point) and to start the trial by pressing a button. After a variable foreperiod (500 ms on half the trials, 1000 ms on other half, selected unpredictably), either the small or the large pre-cue appeared (randomly, but with equal probability). Participants were informed that they did not need to respond in any way to the pre-cue. One hundred fifty ms after the onset of the pre-cue, the character array appeared and was displayed for 50 ms. The pre-cue remained on screen until the end of the trial. Immediately upon removal of the character array (i.e., stimulus onset asynchrony, SOA = 50 ms), one of the three partial-report tones (selected randomly) was sounded. Participants were instructed to report the top, middle, or bottom row of

characters when presented with the high, middle, or low tone, respectively. Participants were allowed to move their eyes following the partial-report tone in order to respond: for each trial, participants wrote the characters they recalled from the cued row in the corresponding row of a 3×4 grid printed on a sheet of paper; they were instructed to guess for any characters they could not recall (Sperling, 1960).

In the placeholders present condition, a placeholder stimulus (# symbol) was presented at each of the 12 locations of the impending alphanumeric characters; these placeholders were present throughout the entire foreperiod and pre-cue period, and all other methods were the same as above.

After participants were given task instructions at the beginning of the session, they next performed a tone-discrimination task in order to confirm that the three partial-report tones were easily discriminated and that participants had learned the tone-to-row mapping: In each trial, a randomly selected high, middle, or low tone was played, and the participant was required to indicate, vocally, the corresponding row (“high,” “mid,” or “low”). The experimenter recorded the participant’s response for each successive tone, and the sequence was discontinued when responses were correct for at least three presentations of each tone. Participants rapidly learned the tone-to-row mapping and performed the task with high accuracy.

Following the tone-discrimination task, participants performed 12 practice and 80 test trials of the task either with or without placeholders, followed by 12 practice and 80 test trials in the other condition. The order of conditions (placeholders absent, placeholders present) was counterbalanced across participants.

2.4. Results and discussion

We estimated the total number of characters available in VSM by multiplying the mean number of characters correctly reported for the cued row by the number of rows in the array (the same procedure used by Sperling, 1960). The primary finding of Experiment 1 was that, with placeholders absent, the estimated number of characters available in VSM was reliably greater with presentation of the small pre-cue than with the large pre-cue ($M(SE) = 7.1, SE = 0.5$ items and $M = 6.6, SE = 0.6$ items, respectively; $F(1, 13) = 4.80, p = 0.047$, partial $\eta^2 = 0.27$; Fig. 1b). These results are consistent with the hypothesis that narrowly focused attention increases the effective capacity of VSM.

With placeholders present, the cue-size effect was eliminated: the estimated capacity with the large pre-cue was 5.6 ($SE = 0.5$) characters, and the estimated capacity with the small pre-cue was 5.2 ($SE = 0.5$) characters; this difference (with a trend in the direction opposite that observed without placeholders) was not reliable ($F(1, 13) = 2.74, p = 0.12$, partial $\eta^2 = 0.17$). Furthermore, while estimated capacity overall was reliably lower with placeholders present than absent ($F(1, 13) = 26.93, p = 0.001$, partial $\eta^2 = 0.67$), there was also a reliable interaction between placeholder condition (present vs. absent) and cue size ($F(1, 13) = 7.08, p = 0.02$, partial $\eta^2 = 0.35$), confirming that the number of characters reported differed in the small and large-cue conditions only when there were no placeholders. These results are consistent with the hypothesis that sustained voluntary attention to stimulus-indexed, task-relevant locations can suppress the automatic rescaling of attention elicited by irrelevant stimuli, at least under some conditions. Moreover, because the simple addition of placeholders eliminated the cue-size effect on estimated memory capacity, these results rule out alternative explanations that the cue-size effect observed in the primary condition arose merely due to physical

display differences confounded with cue size that were present during the exposure of the character array, such as the greater distance between the character array and the border of the large pre-cue, or the greater proportion of the cued region occupied by the character array within the small pre-cue.

At this point it is important to consider two factors that could influence the interpretation of the results. First, whether the 150-ms SOA between the onset of the pre-cue and the onset of the character array was sufficient to allow observers to scale the focus of attention. Several previous studies have provided estimates of the time required to expand or contract the attentional focus. Benso, Turatto, Mascetti, and Umiltà (1998) estimated that it took 33–66 ms to expand the focus of attention to match the size of an abrupt-onset spatial pre-cue. Jefferies and colleagues estimated that the focus of attention could be contracted within 64–103 ms, depending on relevance of the stimuli for the task at hand (Jefferies & Di Lollo, 2009; Jefferies, Enns, & Di Lollo, 2017). In all cases, the estimated time required to rescale the attentional focus is less than the 150-ms SOA between the onset of the cue and the onset of the character array in the present study. In consequence, we can reasonably assume that observers had sufficient time to reflexively rescale the spatial extent of the focus of attention in the present study.

A second point to consider is whether the fixed temporal relationship between the onset of the pre-cue and the onset of the character array 150 ms later introduces an unintended confound by serving as an alerting warning signal. Warning signals enhance performance through top-down mechanisms such as expectancy and through bottom-up processes such as phasic arousal (e.g., Bernstein, Rose, & Ashe, 1970; Coull & Nobre, 1998; Fernandez-Duque & Posner, 1997; Jefferies & Di Lollo, 2017, 2018; Nakayama &

Mackeben, 1989; Schmidt, Gielen, & van den Heuvel, 1984; Spalek & Di Lollo, 2011). Importantly, in the present study the small and the large pre-cue would serve equally as a temporal warning signal, and any expectancy or alerting effects would therefore occur in both the small and large pre-cue conditions. Thus, while we cannot dismiss the possibility that the fixed interval between the cue and the character array resulted in the cue serving as a temporal warning signal, we can be confident that the difference in the effective capacity of VWM when attention is narrowly as opposed to broadly scaled cannot be attributed solely to the warning effect of the cues.

In summary, Experiment 1 indicates that the automatic narrowing or broadening of visuospatial attention in response to an abruptly presented pre-cue influences the amount of stimulus information available in VSM. These results are consistent with the hypothesis that the breadth of attention modulates the effective capacity of VSM. However, because previous studies suggest differential effects of transient, involuntary orienting of attention vs. sustained, voluntary orienting of attention (e.g., Carrasco & Yeshurun, 2009; Prinzmetal, McCool, & Park, 2005), it is possible that important differences likewise exist for involuntary vs. voluntary attentional *scaling* mechanisms. For example, the effect of the breadth of attention on effective memory capacity evidenced in Experiment 1 might be limited only to circumstances eliciting the involuntary capture of attention. In Experiment 2, we tested whether voluntary deployment of attention to a small or large cued region would similarly influence the effective capacity of VSM.

3. Experiment 2

3.1. Participants

Eight participants (M (SE) age = 20.9 (0.6) years; 6 females, 2 males) reporting normal hearing and normal or corrected-to-normal vision were included in the analyses reported below. Participants were paid \$10/hr and were naïve as to the purpose of this experiment.

3.2. *Stimuli and apparatus*

Stimuli and apparatus were the same as in Experiment 1, except that on 4-item trials (described below), two letters and two digits were presented on an imaginary circle either 2.2° or 8.3° in diameter, centered on the fixation stimulus. The four characters were each separated by 90° of arc but at unpredictable locations on the imaginary circle across trials. Also, the pre-cues were gray and had a border thickness of 0.06°. Eye position was monitored by the same method as in Experiment 1.

3.3. *Procedure*

In Experiment 2, *both* the large and small pre-cues were present on every trial (Fig. 2a). Participants were instructed at the beginning of each of the eight blocks of trials (40 trials per block) which one of the pre-cues (large or small) was task-relevant, and that all characters would always appear within the space encompassed by the relevant pre-cue. The relevant pre-cue (large or small) alternated across blocks, with block order counterbalanced across participants. On a random 25% of the trials within each block (Fig. 2a, top row), the 3 × 4 character arrays were presented after a 500-ms or 1000-ms foreperiod, in the same location as in Experiment 1 (i.e., centered within the smaller cue box), followed by the partial-report tone that indicated which row of the array to report. Critically, these trials had physically identical display characteristics across trial blocks, the only distinction being which pre-cue was relevant at the time of presentation.

On the remaining 75% of the trials within each trial block, instead of the 3 x 4 array, four characters appeared after a 500-ms or 1000-ms foreperiod, within the space of the relevant pre-cue, and participants were instructed to report the identities of all four items; the partial-report tone was therefore omitted. These trials were included only to ensure that participants would voluntarily deploy their attention to the region indicated by the relevant pre-cue. In blocks where the relevant pre-cue was small, the four items appeared on an imaginary circle within the space of the small pre-cue (Fig. 2a, middle row). In blocks where the relevant pre-cue was large, the four items appeared on a large imaginary circle, within the space between the borders of the large and small pre-cues (Fig. 2a, bottom row).

Other details of the procedure were the same as in Experiment 1, except that participants performed 160 test trials in each of two sessions conducted less than a week apart.

b

Fig. 2 Schematic illustration of the task in Experiment 2. Both large and small pre-cues were always present throughout the experiment (including the inter-trial interval), but for each block of trials only one of the pre-cues was task-relevant. In a random 25% of trials within each block (*Panel a, top row*), after a foreperiod of 500 or 1000 ms, a 3 x 4 array of alphanumeric characters was presented for 50 ms. Immediately after the offset of the character array, a partial-report tone indicated which row of the characters participants should recall. In the remaining 75% of trials within each block (*Panel a, middle and bottom rows*), four characters appeared on an imaginary circle within the region of the task-relevant pre-cue, and participants were to recall all four characters. Stimuli are not drawn to scale. *Panel b*: Mean estimated capacity (number of characters available in VSM) as a function of the task-relevant pre-cue. Error bars represent the standard error of the cue-size effect

3.4. Results and discussion

Accuracy was relatively high in 4-item trials when the large pre-cue was task-relevant ($M = 94.8\%$, $SE = 1.6\%$ correct). Given the very brief presentation and widely spaced locations of these stimuli, this finding suggests that participants were successful at expanding the focus of attention to encompass the region bounded by the relevant pre-cue even for the large pre-cue. Accuracy was also high in 4-item trials when the small pre-cue was task-relevant ($M = 97.7\%$, $SE = 0.6\%$).

The estimated capacity of VSM was calculated based on participants' performance on the 3×4 character-array trials. Capacity was significantly larger ($M = 6.17$, $SE = 0.54$ items) with a small relevant pre-cue than with a large relevant pre-cue ($M = 5.36$, $SE = 0.57$ items; $F(1, 7) = 25.02$, $p = 0.002$, partial $\eta^2 = 0.78$), as shown in Figure 2b. All eight participants exhibited this cue-size effect by the second session, with the small pre-cue condition producing a mean 23% increase in estimated capacity (range = 5% to 45% across participants) relative to that observed in the large pre-cue condition. These results demonstrate that the voluntary, endogenous rescaling of attention modulates the effective capacity of VSM.

Finally, across experiments, analyses restricted to the placeholders absent condition in Experiment 1 (since there was no placeholder present condition in Experiment 2) indicated that whereas the main effect of cue-size was reliable ($F(1, 20) = 15.98$, $p < 0.001$, partial $\eta^2 = 0.44$), there was not a reliable main effect of experiment ($p > 0.05$), nor was there a reliable interaction between experiment and cue size ($p > 0.05$). These results suggest that exogenous and endogenous rescaling of attention might be

comparable in their influences on the effective capacity of VSM, a possibility that should be studied further.

4. General Discussion

The contents of VSM are fleeting, yet they provide a rich store of detailed visual information, some portion of which can be voluntarily selected for transformation into more durable working memory representations that, in turn, constitute the basis for higher visual cognition.

Attention plays multiple roles in VSM. It serves an important function in the transformation of information from VSM into working memory representations, as evidenced by Sperling (1960) and supported by subsequent findings (e.g., Ruff et al., 2007). The present results suggest that the breadth of attention plays a fundamental role in visual processing by influencing the amount, quality, and/or duration of stimulus information that is encoded into VSM and therefore available for retrieval. It of course was not a foregone conclusion that cue size would affect the effective capacity of VSM in this way, since previous studies that have manipulated several factors expected to aid the discriminability and therefore encoding into memory of stimuli produced negligible effects on the effective capacity of VSM. For example, Sperling (1960) varied the exposure duration of the stimulus array up to 500 ms, but participants showed no systematic changes in the number of letters correctly reported as the exposure duration was increased. Adelson and Jonides (1980) tested whether brighter or higher-contrast stimuli resulted in longer-lasting iconic memory but found that there was virtually no effect on partial-report accuracy. Furthermore, several studies (e.g., Fu et al., 2005; Luo et al., 2001; Wang et al., 2010) have found that cue size reliably affects both visual search

performance and the amplitudes of event-related potentials that reflect the earliest stage of visual processing of target stimuli modulated by attention. Whether the influence of attention scaling continues downstream to mechanisms of encoding and consolidation into working memory requires further study and more broadly motivates an interesting and important investigation into the extent to which attention scaling influences information processing at higher cognitive levels.

Our results dovetail with previous evidence that the efficiency and quality of visual processing is inversely related to the breadth of attention (e.g., Castiello & Umiltà, 1990; Eriksen & St. James, 1986; Eriksen & Yeh, 1985; Jefferies, Gmeindl, & Yantis, 2014; Maringelli & Umiltà, 1998; Turatto et al., 2000); our findings reveal that this relationship impacts VSM as well.

In Experiment 1, an unpredictably large or small pre-cue appeared abruptly prior to an alphanumeric stimulus display and triggered a reflexive rescaling of the focus of attention (Maringelli & Umiltà, 1998; Turatto et al., 2000; Fu et al., 2005; Luo et al., 2001; Wang et al., 2010). The ensuing cue-size effect on recall performance indicates that the automatic narrowing or broadening of visuospatial attention modulates the quantity and/or quality of information from the stimulus display represented in VSM. Moreover, when participants were presented with task-relevant placeholder stimuli intended to prevent the automatic rescaling of attention to the pre-cues, the cue-size effect was eliminated. Previous research has demonstrated that the involuntary *orienting* of attention to abrupt onsets can be avoided when participants engage highly focused attention elsewhere (e.g., Yantis & Jonides, 1990). The current results suggest that the automatic *rescaling* of attention likewise can be avoided when attention is voluntarily

anchored to stimuli that index task-relevant locations. This novel finding accords with previous suggestions that attention focusing is subject to only “weakly voluntary” control, and that people are unable to control the breadth of attention in the absence of perceptual objects to which attention can be anchored (Jefferies & Di Lollo, 2015, 2017; Turatto et al., 2000).

In Experiment 2, we tested whether the cue-size effect observed in Experiment 1 would also arise when attention is voluntarily narrowed or broadened, in the absence of abrupt cues, given that previous investigations have revealed different effects of involuntary and voluntary orienting of attention in other domains (e.g., Carrasco & Yeshurun, 2009; Prinzmetal et al., 2005). We found robust evidence for cue-size effects during the voluntary deployment of attention for all participants, with a mean increase in estimated VSM capacity of 23%. These results show that both voluntary, endogenous rescaling of attention and involuntary, automatic rescaling of attention to abruptly presented objects modulate the effective capacity of VSM.

The tradeoff between the breadth of attention and the quality of visual processing reported in previous studies suggests that the breadth of attention influences the *quality* of the stimulus properties encoded into VSM. For example, the decreased attention allocated to any given stimulus appearing within a relatively broad focus of attention (e.g., Erikson & St. James, 1986) may result in a less precise, less faithful representation of the stimulus. Carrasco and colleagues found that narrowly focused attention enhances the spatial resolution of visual processing (Carrasco & Yeshurun, 2009; Montagna, Pestilli, & Carrasco, 2008; Yeshurun & Carrasco, 2008; Yeshurun, Montagna, & Carrasco, 2008). For the partial-report task employed here, it is therefore consistent that

performance should benefit from the increased resolution of visuospatial processing conferred by a narrow focus of attention, and, conversely, suffer due to the decreased spatial resolution accompanying a broad focus of attention. Reduced precision of the individual stimulus representations encoded into VSM not only would make the rapidly fading trace of a given stimulus more difficult to identify, but also would reduce the discriminability among multiple, concurrent stimulus representations (cf. Bays & Husain, 2008). Both of these effects of reduced spatial resolution would reduce the stimulus information available to be encoded and consolidated into working memory representations.¹

The breadth of attention might also influence the *duration* of information in VSM. Once considered a unitary phenomenon, VSM has been partitioned into theoretical substages or mechanisms based in part on their sensitivity to various stimulus manipulations (e.g., Coltheart, 1980; Di Lollo, 1980). Visible persistence and informational persistence, for example, are differentially affected by stimulus energy and stimulus duration (Irwin & Thomas, 2008). Of particular relevance here are findings that attention decreases the temporal resolution of visual information processing (Yeshurun & Levy, 2003), perhaps by increasing the duration of visible persistence (Visser & Enns, 2001), an early substage of VSM. If it is the case that narrowly focused attention increases the duration of visible persistence and/or informational persistence, then more

---

¹ It is worth noting that the present estimates of effective memory capacity, although expressed in terms of number of items or characters (Sperling, 1960), may simply reflect the total amount of stimulus information (e.g., bits) stored in VSM. The improved quality of visual information (e.g., increased detail) resulting from a narrow focus of attention should nevertheless result in a greater *number* of the display items being accurately reported.

information could be read out from VSM before it is lost, resulting in a greater proportion of the visual display being accessible for transfer into working memory and for subsequent report.

These potential explanations are not mutually exclusive; the breadth of attention could play a critical role in VSM by modulating both the quality and the duration of stimulus representations. The considerable influence exerted by both voluntary and involuntary rescaling mechanisms of attention on the amount of information available in VSM raises interesting new questions about the extent to which these influences generalize across information domains, sensory modalities, and time scales of memory.

Dataset availability statement

The datasets generated during the current study are available from the corresponding author upon reasonable request.

Compliance with Ethical Standards

This study was funded by the National Institute on Aging (T32-AG027668), a New Researcher Grant from the Menzies Health Institute Queensland, and the National Institutes of Health (R01-DA013165). All authors declare that they have no conflict of interest. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

- Adelson, E. H., & Jonides, J. (1980). The psychophysics of iconic storage. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 486-493.
- Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, 40, 309-328.
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, 321, 851-854.
- Benso, F., Turatto, M., Mascetti, G., & Umiltà, C. (1998). The time course of attentional focusing. *European Journal of Cognitive Psychology*, 10, 373-388.
- Bernstein, I. H., Rose, R., & Ashe, V. (1970). Preparatory state effects in intersensory facilitation. *Psychonomic Science*, 19, 113-114.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433-436.
- Carrasco, M., & Yeshurun, Y. (2009). Covert attention effects on spatial resolution. In N. Srinivasan (Ed.), *Progress in Brain Research*, Vol. 176 (pp. 65-86). Amsterdam: Elsevier.
- Castiello, U. & Umiltà, C. (1990). Size of the attentional focus and efficiency of processing. *Acta Psychologica*, 73, 195-209.
- Coltheart, M. (1980). The persistence of vision. *Philosophical Transactions of the Royal Society London B*, 290, 57-69.
- Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience*, 18, 7426-7435.

- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-185.
- Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General*, 109, 75-97.
- Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, 40, 225-240.
- Eriksen, C. W., & Yeh, Y.-Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 583-597.
- Eriksen, C. W., & Yeh, Y.-Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 583-597.
- Fernandez-Duque, D., & Posner, M. I. (1997). Relating the mechanisms of orienting and alerting. *Neuropsychologia*, 35, 477-486.
- Fu, S., Caggiano, D. M., Greenwood, P. M., & Parasuraman, R. (2005). Event-related potentials reveal dissociable mechanisms for orienting and focusing visuospatial attention. *Cognitive Brain Research*, 23, 341-353.
- Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: Individual differences in visual attention control. *Journal of Experimental Psychology: General*, 136, 217-240.
- Irwin, D. E., & Thomas, L. E. (2008). Visual Sensory Memory. In S. J. Luck and A. Hollingworth (Eds.), *Visual Memory* (pp. 9-41). New York: Oxford University Press.

- Jefferies, L. N., & Di Lollo, V. (2009). Linear changes in the spatial extent of the focus of attention across time. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 1020.
- Jefferies, L. N., & Di Lollo, V. (2015). When can spatial attention be deployed in the form of an annulus? *Attention, Perception, & Psychophysics*, 77(2), 413-422.
- Jefferies, L. N., & Di Lollo, V. (2017). Deployment of spatial attention to a structural framework: exogenous (alerting) and endogenous (goal-directed) factors. *Attention, Perception, & Psychophysics*, 79(7), 1933-1944.
- Jefferies, L. N., Enns, J. T., & Di Lollo, V. (2017). The exogenous and endogenous control of attentional focusing. *Psychological Research*, 1-18.
- Jefferies, L.N., & Di Lollo, V. (in press). Sudden events change old visual objects into new ones: A possible role for phasic activation of Locus Coeruleus. *Psychological Science*.
- Jefferies, L. N., Gmeindl, L., & Yantis, S. (2014). Attending to illusory differences in object size. *Attention, Perception, & Psychophysics*, 76(5), 1393-1402.
- Long, G. M. (1980). Iconic memory: A review and critique of the study of short-term visual storage. *Psychological Bulletin*, 88, 785-820.
- Luo, Y.-J., Greenwood, P. M., & Parasuraman, R. (2001). Dynamics of the spatial scale of visual attention revealed by brain event-related potentials. *Cognitive Brain Research*, 12, 371-381.
- Maringelli, F. & Umiltà, C. (1998). The control of the attentional focus. *European Journal of Cognitive Psychology*, 10, 225-246.

- Montagna, B., Pestilli, F., & Carrasco, M. (2008). Attention trades off spatial acuity. *Vision Research*, 49, 735–745.
- Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, 29, 1631-1647
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437-442.
- Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: Reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General*, 134, 73-92.
- Ruff, C. C., Kristjánsson, Á., & Driver, J. (2007). Readout from iconic memory and selective spatial attention involve similar neural processes. *Psychological Science*, 18, 901-909.
- Schmidt, R. A., Gielen, S. C., & Van Den Heuvel, P. J. (1984). The locus of intersensory facilitation of reaction time, *Acta Psychologica*, 57, 145-164.
- Spalek, T. M., & Di Lollo, V. (2011). Alerting enhances target identification but does not affect the magnitude of the attentional blink. *Attention, Perception, & Psychophysics*, 73, 405-419.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74, 1-29.
- Turatto, M., Benso, F., Facoetti, A., Galfano, G., Mascetti, G. G., & Umiltà, C. (2000). Automatic and voluntary focusing of attention. *Perception & Psychophysics*, 62, 935-952.

- Visser, T. A. W., & Enns, J. T. (2001). The role of attention in temporal integration. *Perception, 30*, 135-145.
- Wang, Y., Wu, J., Fu, S., & Luo, Y. (2010). Orienting and focusing in voluntary and involuntary visuospatial attention conditions: An event-related potential study. *Journal of Psychophysiology, 24*(3), 198–209.
- Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 121-134.
- Yeshurun, Y., & Carrasco, M. (2008). The effects of transient attention on spatial resolution and the size of the attentional cue. *Perception & Psychophysics, 70*, 104-113.
- Yeshurun, Y., & Levy, L. (2003). Transient spatial attention degrades temporal resolution. *Psychological Science, 14*, 225-231.
- Yeshurun, Y., Montagna, B., & Carrasco, M. (2008). On the flexibility of sustained attention and its effects on a texture segmentation task. *Vision Research, 48*, 80-95.