Activating our "Visual RAM" for Multitasking:

Does Working Memory bias Visual Attention with multiple items and human emotions?

Supervisor: Dr. Samson Chota[1]
Student: Manfredi Guarisco[2]

[1] Department of Social and Behavioural Sciences, Psychology and Experimental Psychology, Visual Psychophysics Lab, Utrecht University, Utrecht, Netherlands.
[2] Department of Social and Behavioural Sciences, Applied Cognitive Psychology, Utrecht University, Utrecht, Netherlands.

E-mail: s.chota@uu.nl; m.guarisco@students.uu.nl.

Abstract

Visual Working Memory (VWM) drives and influences our everyday visual search. VWM search templates (temporary mental representations of specific items) can bias our Visual Attention; however, we do not know yet if two search templates can bias visual search at the same time. In this study, we investigate how VWM biases visual attention with multiple colors and human emotions in a computer screen task. We measured microsaccades with an eye tracker to track covert attention. We ran two Rapid Serial Visual Stimulation (RSVP) visual search tasks where we asked participants to memorize two color stimuli in the first experiment, and two human emotional faces in the second. We probed both memorized items during the RSVP stimuli presentation together with distractors. Results show a VWM biasing effect on visual attention with two items, but not time-specific; both items biased attention in the same way when both were probed in the same trials. Furthermore, we observed that the attentional bias for emotional faces is stronger for only one template when both were probed in the same trials. Thus, we concluded that two VWM search templates (i.e., colors) can bias attention at the same time, and that there is a stronger bias for only one template after memorizing two emotional faces at the same time.

Keywords: working memory, search templates, multiple items, visual search bias, eye tracker.

1. Introduction. VWM and Visual Attention: a duo making up for our "Visual RAM".

Our everyday lives are filled with multiple visual representations to potentially select, retain, process, and eventually store for a few seconds or permanently. Most of the time, however, we just need to temporarily retain a few of them to carry out our everyday goals.

When it comes to *temporarily store* visual information, Visual Working Memory (VWM) is one of the most important cognitive systems. Evidence from its dissociation with verbal content has been reported by, e.g., Smith et al. (1996), where different locations for blood flows have been observed with PET imaging when spatial location tasks, and verbal-memory tasks have been performed by healthy subjects. VWM holds limited capacity for storing and retaining information (within a short term of time), and then performing mental operations on the content being retained (Gazzaniga et al., 2019a).

When it comes to *select* certain visual information during visual search, Visual Attention is an essential cognitive ability, allowing for the allocation of cognitive resources «among relevant inputs, thoughts, and actions while simultaneously ignoring irrelevant or distracting ones» (Gazzaniga et al., 2019b, p. 277).

These two cognitive processes hold a close relationship between each other. The way VWM interacts with visual attention has been widely investigated in the terms where the information we are visually retaining in our mind (i.e., *what we are thinking about*) can influence the latter selection of items within visual search (i.e., *what we are focusing on*) (Awh & Jonides, 2001; Desimone and Duncan, 1995; Downing, 2000; Sreenivasan et al., 2007). As efficiently synthetized by Olivers (2009, p. 1275): «when observers are looking for something specific, objects matching the target object will involuntarily capture attention» (Folk et al., 1992; Folk et al., 2002; Moores et al., 2003).

That is, while working memory is used to maintain a mental representation of the target item of a visual search task (also known as "attentional set", or *Search Template*), attention will then be biased by perceptual stimuli sharing features with that template (Olivers, 2009; Olivers et al., 2011). But what about visual search with multiple items (and thus, multiple search templates)? In a study by Van Moorseelar et al. (2014), it was observed that when multiple items were memorized, none was formed as a search template that interacted with visual perception; simply put, subjects' visual attention was not biased by multiple VWM search templates. However, two other studies with similar visual search task designs observed no reduction in memory capture after increasing the number of items from one to two (Soto & Humphreys, 2008; Zhang et al., 2011). Van Moorseelar and colleagues (2014) concluded with the possibility for future research to find out why multiple VWM search templates did not bias attention in visual search.

To investigate the interconnections and biasing effects between VWM and visual attention, recording and analysis of microsaccades (that is, subtle eye movements with a < 1° visual angle amplitude) is a useful method that is informative about covert attention allocation (Laubrock et al., 2005; Martinez-Conde et al., 2013; Pastukhov et al., 2010). This kind of extremely small eye movements have the same characteristics as normal saccades, but on a factor of around 50 times smaller (Engbert, 2006); microsaccades can also be observed when participants have been explicitly instructed to keep their eyes on a specific fixation point. Microsaccades also tend to be suppressed prior to the visual search to prepare for the task to be performed (Van Loon et al., 2017), and search template-specific bias has also been observed (i.e., when the memorized item was presented for visual search; e.g., Hollingworth et al., 2013; Schreij et al., 2014).

Rapid Serial Visual Stimulation (RSVP) is a method that has already been tested with microsaccades analysis for VWM search templates (e.g., Olmos-Solis et al., 2017): in a few words, it generally consists of a series of visual stimuli being presented in a series of fast-paced timewindows (e.g., 100ms each; *see* Figure 1) with or without the probing of one or more items being memorized.

In a recent study by Olmos-Solis et al. (2017), it was shown that microsaccades are being biased by specific search templates during RSVP in a visual search task. Specifically, by first presenting two color circles on the top and bottom screen, a directional cue about the matching item location was displayed briefly right after, followed by a series of RSVP timewindows containing two color circles to the left and to the right of the fixation point (at a 180° angle): participants were instructed to keep fixation and ignore the RSVP, to then finally report the orientation of the arrow inside the color they memorized in the visual search display at the end of each trial. Surprisingly, a template-specific (i.e., the memorized color) bias has been observed: a larger saccade amplitude was observed to the matching item right after memorizing the item and just before the visual search task to be performed; in other words, participants' eyes were "efficiently" directing gaze towards the matching item during RSVP. Thus, results from Olmos-Solis and colleagues (2017) study seem to support the idea that the attentional bias by VWM content is also related to the expectation of the visual search to perform. However, the search template to be formed in this experimental design was related to one item at a time; thus, eventual biasing effects related to time coming from two search templates still remain an open question for us: can we look for two items at the same time? How would VWM content bias attention in that case?

Simple items such as colored circles are obviously not the only ones that VWM can retain as search templates; within our everyday lives, it is also likely to encounter complex stimuli such as multi-featured objects, like human faces. Human faces are surely a special category of stimuli for human perception, as their features are processed holystically (i.e., eyes, nose, chin structure and other features are recognized in parallel, rather than individually) (Watson & Robbins, 2014, p. 1, in Eysenck & Keane, 2020). To decide whether a face is familiar or not, only 200ms would be needed (Caharel et al., 2014).

Memory facilitation effects for neutral and emotional human faces stimuli in visual search tasks have been widely explored, specifically about how visual perception behaves when facing the six basic emotions (i.e., happy, surprised, disgusted, fearful, angry, sad). There are some studies suggesting a higher detection rate for angry emotional faces (Eastwood et al., 2001; Fox et al., 2000; Öhman et al., 2001; Pinkham et al., 2010) and happy emotional faces (Williams et al., 2005), with some conflicting evidence from e.g., Calvo et al., 2008, where faces with happy, surprised, and disgusted emotions were detected more effectively and faster than faces with fearful, angry, and sad emotions (compared to neutral faces).

Facial emotions can also facilitate visual search task performances when used as VWM search templates, whereas negative emotions (i.e., angry and fearful) have been associated with higher performance (Gambarota & Sessa, 2019; Jackson et al., 2013; Lee & Cho, 2019; Lingxia et al., 2016), which is probably due to the activation of the amygdala during the encoding and consolidation processes (Lee & Cho, 2019; Phelps, 2004). Even if evidence from previous studies shows that happy and neutral faces stimuli did not differ in VWM memory facilitation (Jackson et al., 2013; Sessa, 2011), facilitation effects for happy faces compared to neutral expressions have been recently found by Lee & Cho, 2019. So far, it is quite hard to find converging conclusions on the matter.

To summarize, it has been widely observed that VWM search templates can bias attention, and thus visual search strategy. However, the biasing effects coming from two search templates at the same time were not explored yet. In this study, we will first investigate if two VWM attentional sets (or search templates) can generally bias visual attention in the presence of multiple colors, and multiple human emotional and neutral faces with a second experiment, by measuring microsaccades during RSVP. We have chosen to address this question by presenting two items to memorize as VWM search templates at the same time in a RSVP visual search task design. Second, we will explore the possibility of only one of the two search templates biasing more attention if both were probed in the same trials.

2. Method and Experimental Design

We have chosen to carry out two separate experiments with an RSVP visual search task design similar to the one used by Olmos-Solis et al. (2017). In the first experiment, we asked participants to memorize two (different) color stimuli. In the second experiment, we presented two (different) emotional or neutral human faces to memorize instead. During the RSVP timewindows we randomly probed the matching items together with distractors, to see if there is indeed a bias in visual attention through subtle gaze behavior.

For both experiments, an EyeLink 1000 Plus (software ver. 5.09) has been used to record participants' monocular gaze behavior at a 1000Hz sample rate in a dimly lit and silent room, with the sole additional presence of the experimenter

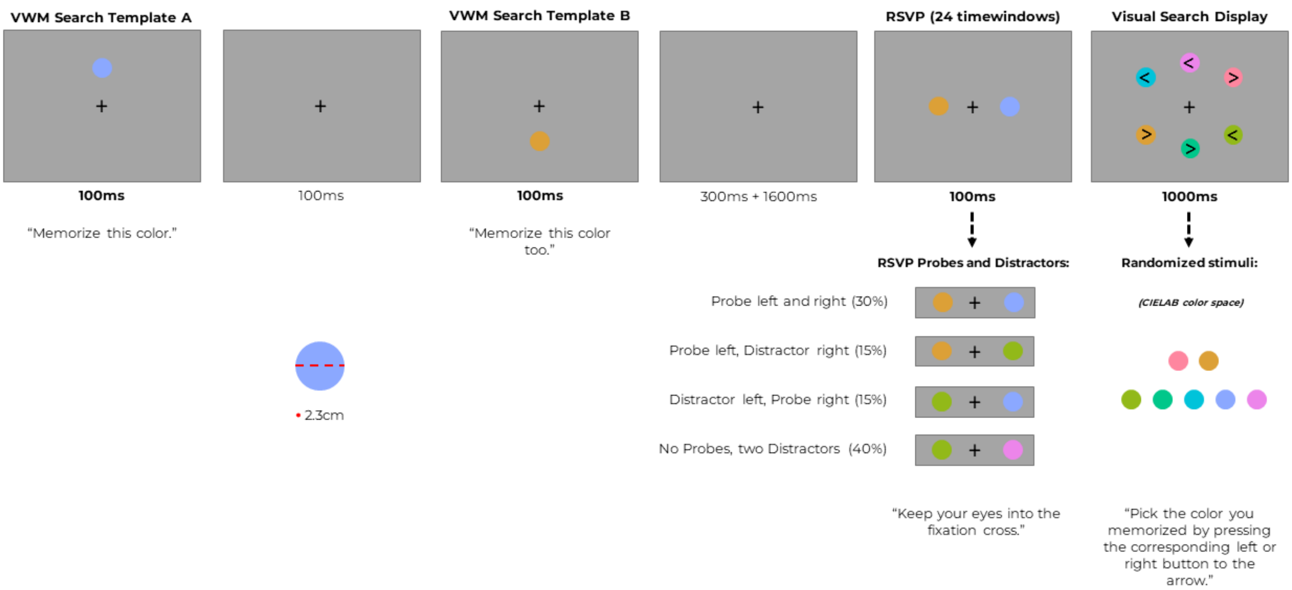

Figure 1. Trial example for Experiment 1. This is a visual search task design with 24 RSVP timewindows (100ms each), which we used to investigate how VWM could bias attention after two search templates (A and B) were being formed at the same time with two different CIELAB colors. A total of six color probes matching to the two colors memorized in the beginning were being displayed with different probabilities to appear during RSVP on each trial. We recorded eye saccades during RSVP to see if there is a biasing effect when the items probed are matching to the memorized colors in the beginning.

during data collection. A total of 8 calibrations have been done for each participant (one calibration and validation every sixty RSVP trials). Every calibration consisted of thirteen circular dotted dark grey points (1cm diameter) with a circular center (0.3cm diameter), randomly changing fixed screen positions during the process.

Experiment 1 and 2 visual search tasks were built with Matlab R2016a (ver. 9.0.0.341360) with Psychtoolbox. Both types of stimuli were presented on a computer display (LCD, 2560 × 1440 resolution, 120Hz refresh rate, 27 inches, 16:9 aspect ratio). An adjustable plastic tower-type support has been used to hold participants' heads in a fixed position (with their chin and forehead resting on it), 78cm away from the display. The eye tracker was positioned to be in line under the computer display.

2.1 Experiment 1: RSVP visual search task design with two color stimuli

Participants

31 participants aged 18 to 32 (*M*: 24.30; *SD*: 3.24) with normal or corrected-to-normal vision took part in the first experiment. The participant pool consisted of 22 females and 9 males, who had already attended 0 to 4 times other experiments with an eye tracker (*M*: 0.96; *SD*: 1.06; Not stated: 4); 27 of them were right-handed, while 4 were left-handed.

Each participant was asked to remove any make-up (e.g., mascara, eyeliner) and/or to wear contact lenses if needed before the experiment. After consent forms were signed, participants were explained the visual search task to perform with a 2-minute PowerPoint presentation and 10 trials as a pre-experiment training. When the experiment was completed, participants were compensated with 12€ or study credits. Participants were further motivated by the contest of a free Pasta alla Carbonara (with guanciale and pecorino) for the best performer.

Experimental Design

Stimuli presentation consited in a central fixation black cross (0.3cm), a 1° view-angle diameter color circles (2.3cm diameter), in a grey background (RGB 80, 80, 80). We used the CIELAB color palette for the items to be presented (hex-code for red: #FE869F; yellow: #DCA037; green: #92B919; teal: #00C78A; blue: #00C3DA; lilac: #8BA8FF; pink: #EC85EA), to avoid the possibility of participants verbalizing the colors to memorize during the task.

We asked the participants to memorize the first two colors (randomized, 100ms each; *see* Figure 1 for further details) while looking at the fixation cross, in order to let VWM form two search templates at the same time. We also asked participants not to verbalize any of the colors. After an interval of 300ms + 1600ms (with no items being displayed other than

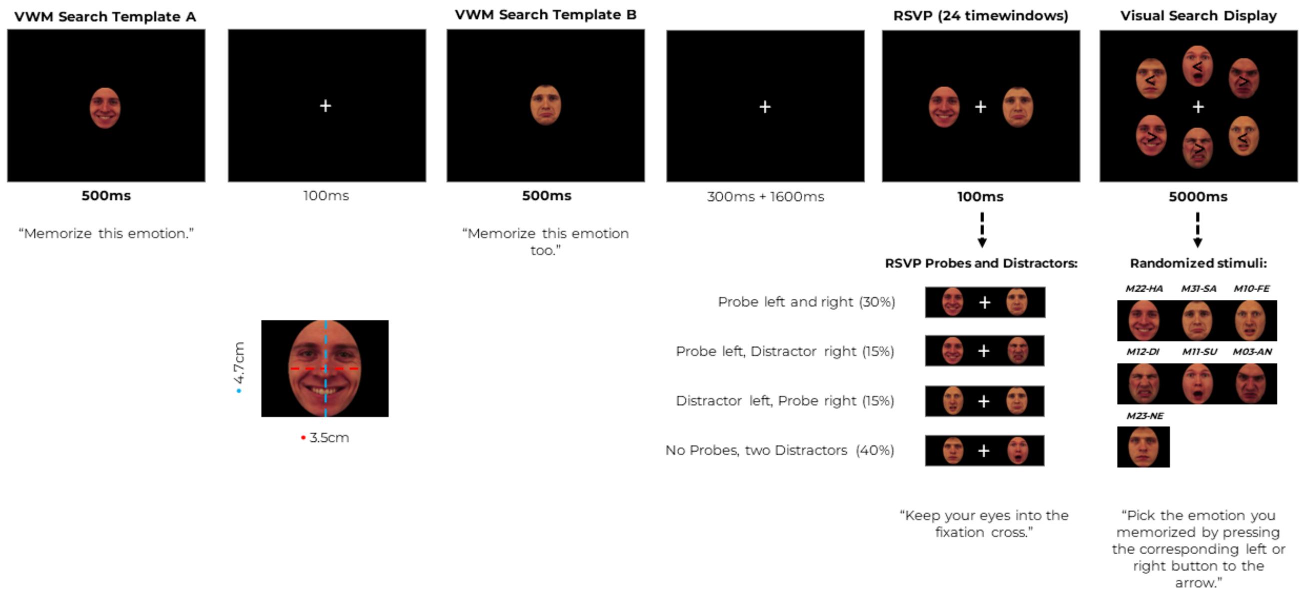

Figure 2. Trial example for Experiment 2. The design is almost identical to the one we used for Experiment 1, but with human emotional and/or neutral faces stimuli and slightly different timings (*see* ch. 2.2). A total of six human emotions/neutrality probes matching to the two emotions/neutrality memorized in the beginning were displayed during RSVP in each trial (KDEF-I: M22-HA, M31-SA, M10-FE, M12-DI, M11-SU, M03-AN, M23-NE). Once again, we recorded eye saccades during RSVP to see if there is a biasing effect when the items probed are matching to the emotional faces memorized in the beginning.

the fixation cross), we asked participants to keep looking at the fixation cross while distractor and probe color stimuli were being presented to the left and to the right side of the cross, throughout all the RSVP (24 timewindows of 100ms each); in each trial, within the RSVP timewindows, we randomized the presentation of matching probes (that is, matching to the two colors memorized in the beginning) by weighting their probabilities to appear as it follows:

- Matching probe to the left and right: 30%;
- Matching probe to the left and random distractor to the right: 15%;
- Matching probe to the right and random distractor to the left: 15%;
- Random distractors to the left and right: 40%.

In the end of the trial, we let participants move their sight away from the fixation point to perform a visual search task where they had to report the orientation of the arrow inside the color they memorized in the beginning, from a total of 6 colors being displayed. Participants had 1 second to answer by pressing the left or right button on the matching arrow's direction (inside the colored circles, randomized). This way, participants were incentivized to move their sight during the visual search display, keeping their eyes in a saccade-ready state throughout all the trials.

If the given answer was correct, the central fixation point would have been turned green. A performance meter was displaying participants' own correct answer rate every 20 trials. Two short breaks of 4 minutes each were taken every 180 trials. Each experiment session consisted of an individual dataset of 480 trials (*see* Figure 1 for an Experiment 1 trial example).

2.2 Experiment 2: RSVP visual search task design with two emotional human faces stimuli

Participants

15 participants aged 19 to 26 (*M*: 22.13; *SD*: 2.26) with normal or corrected-to-normal vision took part in the experiment. The participant pool consisted of 10 females and 5 males, who had already attended 0 to 2 times other experiments with an eye tracker (*M*: 0.46; *SD*: 0.67; Not stated: 2); 12 of them were right-handed, and 3 of them were left-handed.

Preparation, consent, explanation, training, and compensation procedures were identical to the ones we carried out during Experiment 1. Participants were further motivated by the contest of a free Pasta alla Carbonara for the best performer here as well.

Experimental Design

The second RSVP visual search task design was identical to the one of Experiment 1, except for the type of stimuli being presented, the background color of the tasks, the first two memory items presentation time and position, and the Visual Search Display answering timewindow. We used the KDEF-I male faces database as the items to be presented (Lundqvist et al., 1998).

Here, stimuli presentation consisted of a central fixation white cross (0.3cm), emotional and/or neutral human faces stimuli (JPG 100%; 608 × 464 resolution, 96 dpi; max. face length: 3.5cm, max. face height: 4.7cm), in a black background (RGB 0, 0, 0). The items to be memorized were the first two emotions (or neutrality) being presented at the center of the display (i.e., two randomized stimuli of neutral, happy, surprised, disgusted, fearful, angry, or sad faces) for 500ms each, with a brief interlude of 100ms. The probability weights of the six matching probes to appear within the RSVP timewindows were arranged exactly as in Experiment 1. At the end of the trial, we granted participants 4 more seconds to answer, compared to Experiment 1 (*see* Figure 2).

2.3 Data analysis

Data analysis for Experiments 1 and 2 was carried out with Matlab R2022b (ver. 9.13.0.2126072) with Statistics and Machine Learning Toolbox, Jamovi (ver. 2.3.21, 2022, based on R, ver. 4.1, 2021), and Microsoft Excel (2023) for graphical results plotting. Data analysis has been processed separately for each experiment.

The eye tracking analysis included only trials with correct answers. We have analyzed microsaccades during the RSVP timewindows (24, 100ms each) by grouping four separated conditions into four larger timewindows of 600ms, allowing us to distinguish between early (0-500ms), middle (600-1100ms and 1200-1700ms), and late time conditions (1800-2300ms), related to the appearance of the visual search display, similar to what has been done by Olmos-Solis and colleagues (2017; *see* "Oculomotor measures reveal the temporal dynamics of preparing for search" for further details).

First, we hypothesized that microsaccades would be biased towards the left or right if the items did match VWM content during RSVP; we also expected that this bias would differ within different time conditions, as it did in Olmos-Solis et al.

(2019) for single-template visual search (*see* Figure 3). We tested our first hypothesis with a repeated measures ANOVA on the averaged microsaccades amplitude during the four RSVP timewindows when memorized items were matching to the left or to the right, for each participant (H_1).

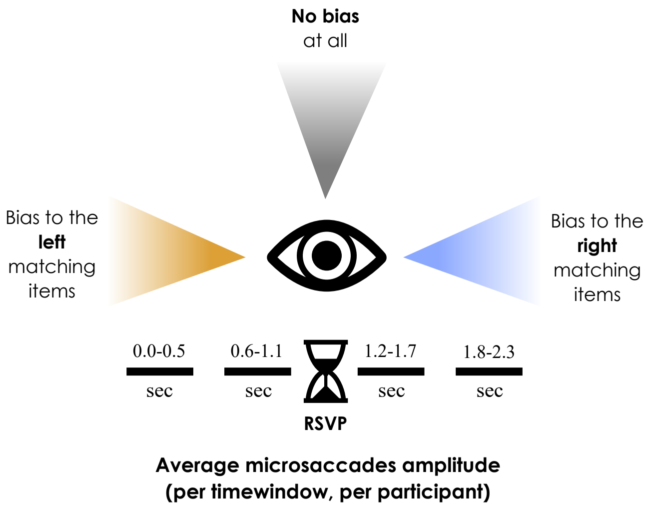

Figure 3. Expectations for our first hypothesis (H_1). If the memorized item is matching to the *left* during RSVP, we expect participants to produce microsaccades to the left (thus, recording negative values in terms of horizontal amplitude coordinates); on the other hand, if the memorized item is matching to the *right*, we expect participants to produce microsaccades to the right, thus recording positive coordinate values. However, if participants do not produce any microsaccades at all when memorized items are matching to the left or right, coordinate values will then be equal to zero.

We also questioned if only one VWM Search Template could bias microsaccades more, when both were probed in the same trials (i.e., attentional bias only to the item memorized first or only to the item memorized second).

We tested this hypothesis by taking into account only trials in which both items A and B were probed during RSVP. We then considered the left and right microsaccades amplitude for item A and B under such conditions, and transformed the leftward amplitude into a positive value by multiplying it by -1; finally, we subtracted the amplitudes of both items between each other's, to obtain a value that we called *attentional difference* for each participant. If this value turned out to be highly positive or negative, it would indicate that only one item is biasing more attention; if the attentional difference value did approach zero, then both items A and B did bias attention in the same way. To test the significance of this second hypothesis, we ran a two-tailed one-sample t-test on the attentional difference for each participant under such conditions (H_2).

We assumed a 95% confidence interval for significance testing for all of our hypotheses. For both experiments, participants that scored less than 55% on their given correct answers were excluded from data analysis (*N*: 6 for Experiment 1). We also had to exclude individual datasets which turned out to include too many missing answers on the visual search task (*N*: 1 for Experiment 1), and individual datasets with more than 2.20° max. error in more than 2 re-calibrations (*N*: 1 for Experiment 1).

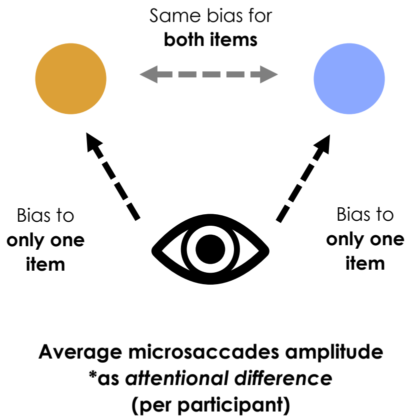

Figure 4. Expectations for our second hypothesis (H_2). If only one of the memorized items (A or B) is biasing more attention, then the *attentional difference* between the microsaccades amplitude when the Search Template A or B were matching (when both appeared in the same trial) should be different from zero. Otherwise, if both items biased attention with the same strength, we would expect the attentional difference to be near or equal to zero.

Ultimately, we took into account 24 participants' datasets for Experiment 1, and 15 for Experiment 2.

3. Results

3.1 Experiment 1: multiple Visual Working Memory search templates do bias attention

We first ran a performance analysis to check the average proportion of correct answers for the visual search task with CIELAB color stimuli (*see* Figure 5). Results showed us how participants performed when the item was memorized first (as VWM Search Template A; *M*: 0.6941; *SD*: 0.0758) and second (VWM Search Template B; *M*: 0.6829; *SD*: 0.0713).

The averaged proportions of correct answers between different colors when being presented as Search Template A and Search Template B as for CIELAB red (as VWM Search Template item A: 0.708; as VWM Search Template item B: 0.711), yellow (A: 0.745; B: 0.710), green (A: 0.720; B: 0.702), teal (A: 0.734; B: 0.695), blue (A: 0.690; B: 0.691), lilac (A: 0.602; B: 0.604) and pink (A: 0.648; B: 0.668) are shown in Figure 5. The widest difference in performance was between yellow and lilac (A: 14.3%; B: 10.6%).

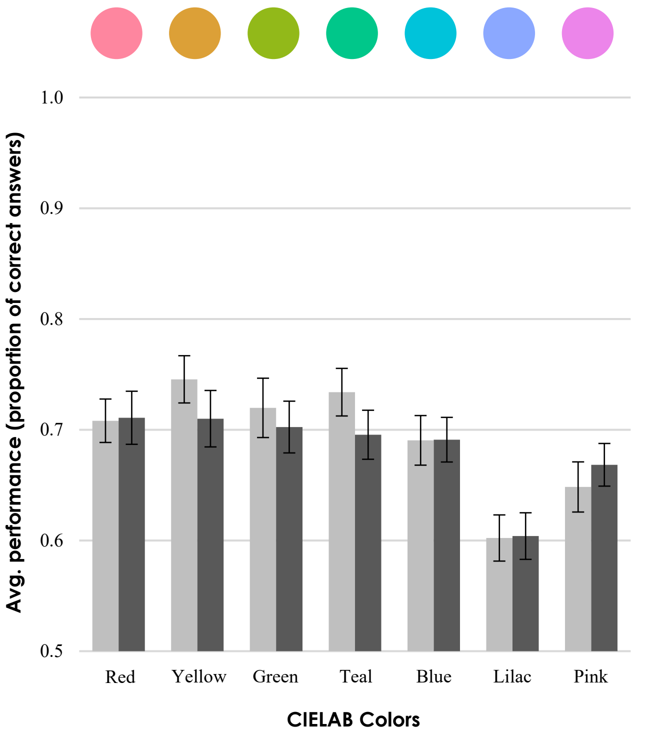

Figure 5. Averaged performances for all participants with different CIELAB colors in the visual search task of Experiment 1. Light grey bars are for performances when the color appeared first (thus, as VWM Search Template *A*), while dark grey bars indicate performances when the color appeared last (as VWM Search Template *B*). Error bars mark their Standard Errors.

We first questioned if two VWM search templates formed at the same time could bias visual attention (thus, eye saccades). To assess this possibility, we took into account the microsaccades amplitude during the 24 RSVP timewindows (100ms each) in trials where a correct answer was given, by then compressing the timewindows into 4 of 600ms each; finally, we averaged that amplitude for each participant every 600ms, by analyzing the two conditions where memorized items were matching to the *left*, and where they were matching to the *right*. Here, we had to remove *N*: 3 subjects due to the lack of microsaccadic production in one late (1800-2300ms) and two early (0-500ms) RSVP timewindows under such conditions.

We ran a repeated measures ANOVA to compare the biasing effect overtime (RSVP timewindows, as our independent variable) on memorized items matching to the *left* and to the *right* (average microsaccades amplitude, as our dependent variable). We found a statistically significant difference within subjects' means on conditions where items were probed to the left, and to the right [H_1: F (df = 1, 20) = 5.282, p = 0.032]. We did not find any main effect in RSVP timewindows [H_1: F (df = 3, 60) = 0.815, p = 0.490], nor in the interaction within left and right matching probes and RSVP timewindows [H_1: F (df = 3, 60) = 1.510, p = 0.221]. Later, Tukey's HSD post-hoc analysis mirrored the effect on the difference within left and right matching condition means (df = 20; t = -2.30; p_{tukey} = 0.032), as well as the absence of an effect in the interaction of the 4 timewindows with the attentional bias. As we found that microsaccades were indeed biased to the left or to the right when the matching items were probed during RSVP, we rejected our first null hypothesis; in other words, these results show that multiple VWM content *does* bias visual attention, even if this bias is not time-specific.

By plotting the average microsaccades amplitude of all participants during RSVP timewindows in all (correct) trials, we observed that microsaccades direction seems to convey to the right during the first 600ms, with its highest amplitude during the late condition (*M*: 0.158; *SE*: 0.163); within the mid RSVP timewindows (600-1700ms), we observed a faint microsaccadic suppression when the item was matching to the right (*see* Figure 6, lilac bars). On the other hand, we noticed a slightly flatter amplitude when the item was matching to the left, yet more constant throughout all the RSVP timewindows: the bias seems to get gradually stronger, with its highest peak as the visual search display is approaching (*M*: -0.121; SE: 0.158; 1800-2300ms; *see* Figure 6, yellow bars).

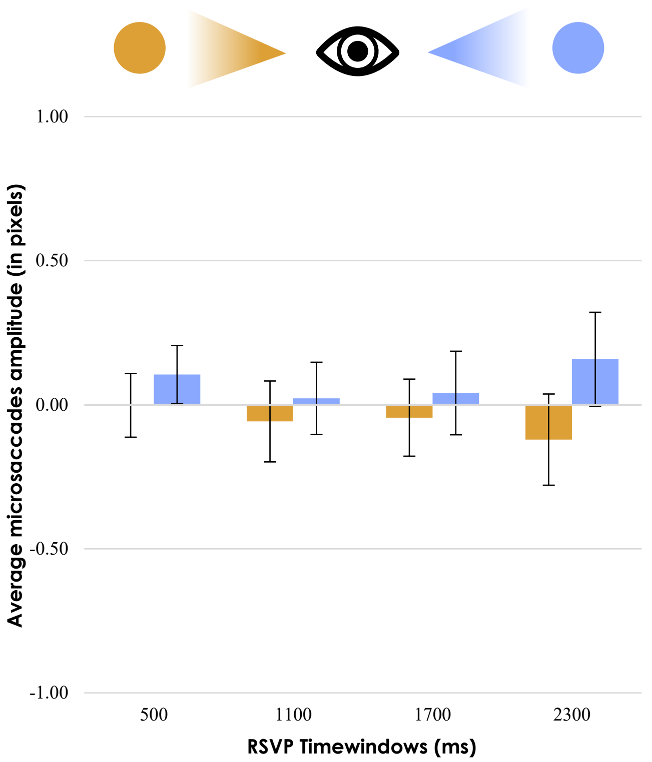

Figure 6. Average microsaccades amplitude during RSVP (compressed into 4 timewindows of 600ms) over all participants in Experiment 1: yellow bars show the average amplitude when the memorized color (one of the randomized shades from the CIELAB color palette) was matching to the *left*, while lilac bars stand for their averaged amplitude when the item was matching to the *right*. Error bars indicate the Standard Error of the mean.

We then questioned if only one VWM item biased *more* microsaccades, when both were probed in the same trials. To answer our second hypothesis, we took trials where both Search Templates A and B were probed during RSVP; then, we computed the average difference in microsaccades amplitude to the left (transformed into a positive value) and to the right between both items, to assess the possibility that one specific template biased more attention, or that both did in the same way. In this case, we had to remove *N*: 2 participants because of the absence of microsaccades production in such trials combination. We ran a two-tailed one-sample t-test, which resulted in a non-significant difference for one specific item (A or B) to bias more microsaccades (H_2: $df = 21$; $t = 1.605$; $p = 0.689$). Thus, we refused to reject our second null hypothesis, meaning that both items biased attention in the same way when probed in the same trials.

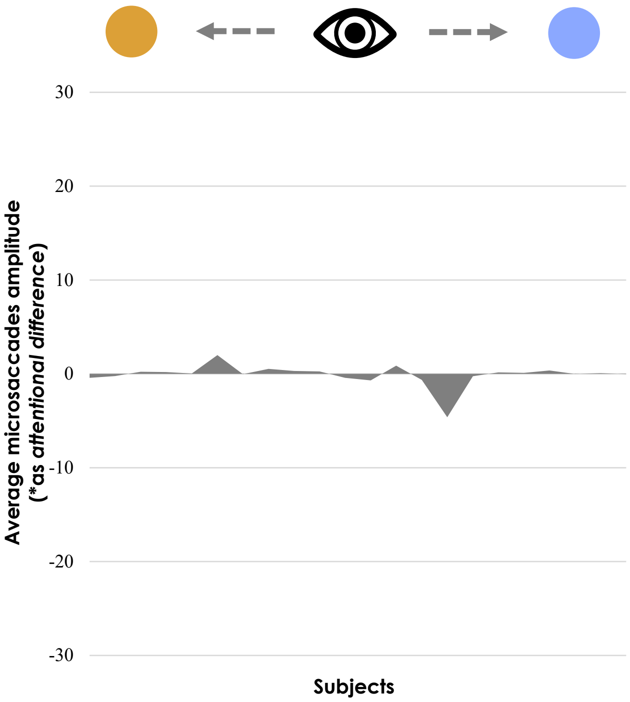

Figure 7. Distribution of the average microsaccades amplitude (*as attentional difference) during RSVP for each participant, in Experiment 1. Participants are distributed horizontally (*N*: 22), while the grey area indicates how strongly their attention was being biased for only one item. Negative and positive values indicate that only one of the two probed colors did bias more attention; but here, values approaching 0 suggest that there is no difference at all ($p = 0.689$).

3.2 Experiment 2: a prioritized, leftward grin

The performance analysis for the second experiment shows an overall higher score compared to Experiment 1. Participants performance for Search Template A (*M*: 0.8861; *SD*: 0.0945) and B (*M*: 0.9072; *SD*: 0.0699) are shown in Figure 8.

Even if some participants reported to memorize happy faces (*N*: 6) and sad faces (*N*: 5) the best, but fearful faces the worst (*N*: 5), the averaged proportions of correct answers between happy (as VWM Search Template item A: 0.921; as VWM Search Template item B: 0.907), sad (A: 0.894; B: 0.904), fear (A: 0.846; B: 0.893), disgust (A: 0.883; B: 0.926), surprise (A: 0.894; B: 0.914), angry (A: 0.893; B: 0.892) and neutral (A: 0.874; B: 0.913) turned out to be quite similar. Several subjects (*N*: 5) also pointed out that focusing on the mouth and eyebrows area of the faces stimuli granted them a better performance in memorizing and later choosing the right emotions to score a correct answer.

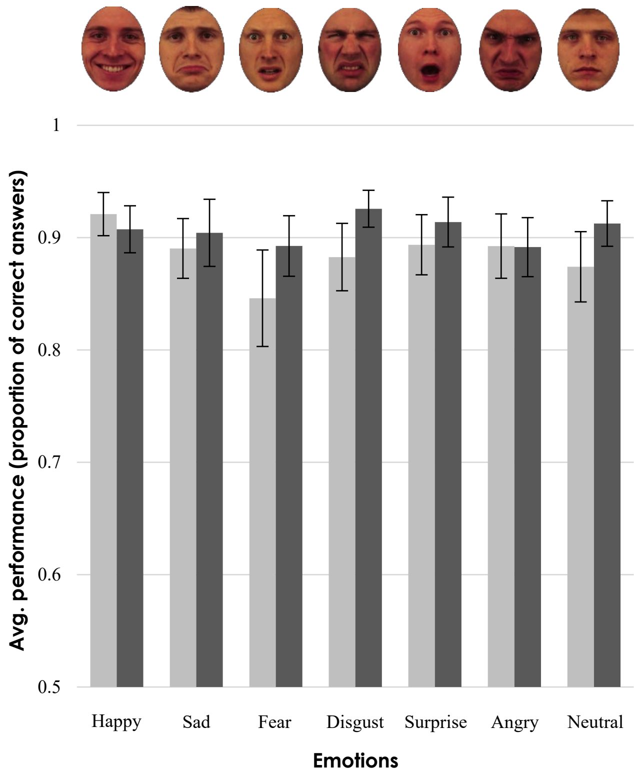

Figure 8. Averaged performances across all participants between different emotions in the visual search task of Experiment 2. Light grey bars are for performances when the emotion appeared first (thus, as Search Template *A*), while dark grey bars indicate performances when the face appeared the last (as Search Template *B*). Error bars stand for their Standard Errors.

We proceeded with the same analysis logic as Experiment 1, and ran an attentional bias analysis, again measured as the individual average microsaccades amplitude for each participant, for each RSVP timewindow compressed into 4 of 600ms each, by taking correct trials only. Once again, we performed a repeated measures ANOVA to compare the biasing effect over time on memorized emotions matching to the *left* and to the *right*. Results are showing a within subjects non-significant difference on conditions where items were probed to the left, and to the right [H_1: F (df = 1, 14) = 0.194, p = 0.666], with a similar outcome within the four RSVP timewindows [H_1: F (df = 3, 42) = 2.065, p = 0.119]. Surprisingly, we found a significant interaction of left and right matching probes with RSVP timewindows [H_1: F (df = 3, 42) = 3.300, p = 0.029]. However, after computing a post-hoc Tukey's test, we did not find any significance in the interaction within subjects throughout all combinations amid dependent and independent variables. We also did not find any other effect, globally (within left and right matching items conditions) nor time-specific (RSVP timewindows). Ultimately, we did not find evidence for multiple emotional faces to bias eye movements when used as VWM search templates; therefore, we refused to reject our first null hypothesis in this case.

We also observed that the distribution of averaged microsaccades amplitude for all participants over RSVP does not reveal the same patterns we retrieved in Experiment 1. In fact, microsaccadic production seemed to convey to the left quite constantly throughout both conditions (that is, when faces were matching to the *left*, *and* to the *right*), with their amplitude spike between 1200ms and 1700ms (when matching to the left, *M*: -13.219; *SE*: 18.874; when matching to the right, *M*: -10.448; *SE*: 18.105).

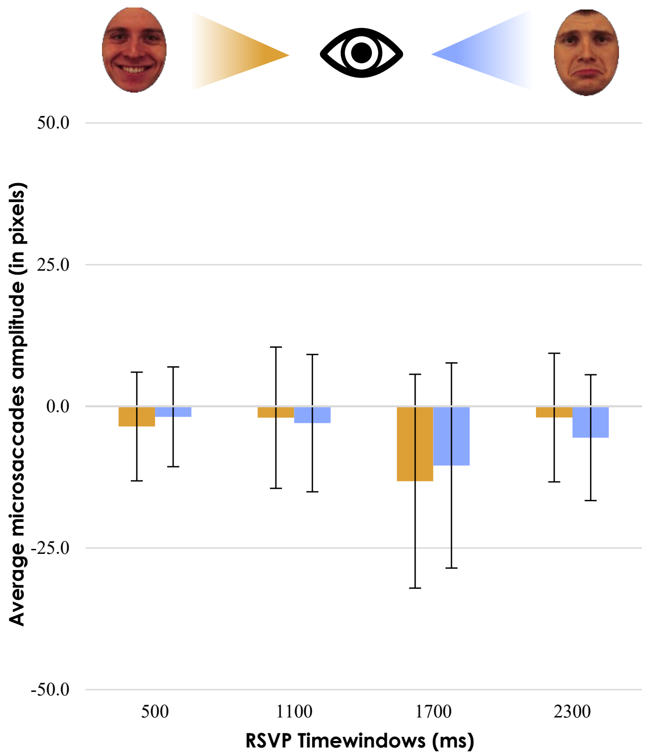

Figure 9. Average microsaccades amplitude during RSVP over all participants in Experiment 2. Once again, yellow bars are showing the average amplitude when the memorized emotional face was matching on the *left*; and lilac bars when they were matching on the *right*. Error bars indicate their S.Es.

Moving to our second hypothesis, and basing the calculations on the same logic we pursued for H_2 in Experiment 1, results from a two-tailed one-sample t-test have shown a significant difference for only one specific item (A or B) to bias more microsaccades (H_2: df = 14; t = 2.656; p = 0.019). Therefore, we rejected our second null hypothesis, meaning that eye movements are mostly directed to only one memorized emotional face at a time, even if both were probed in the same trials.

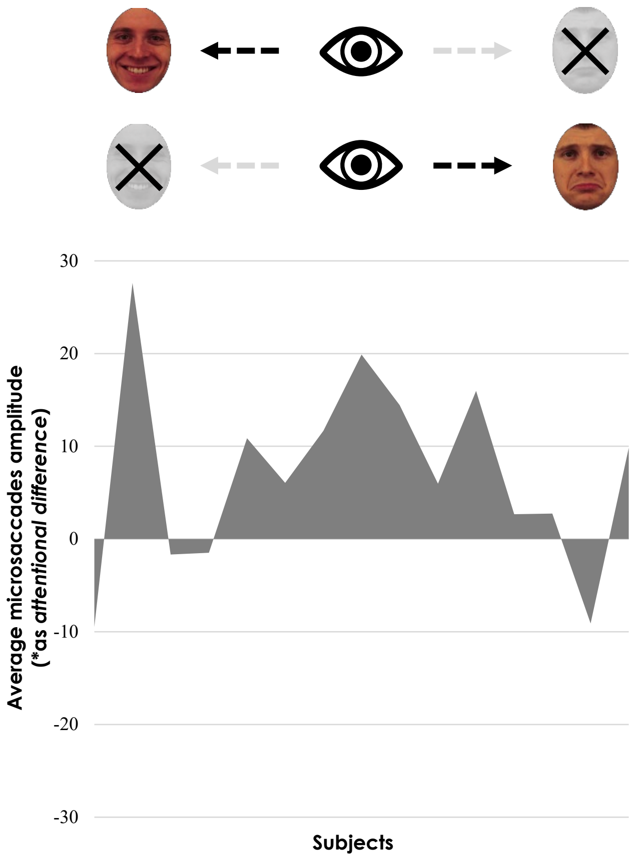

Figure 10. Distribution of the averaged attentional difference during RSVP for each participant, in Experiment 2 (*N*: 15). Participants are distributed horizontally. The grey area indicates how stronger the attentional bias was for only one item (*A* or *B*). Here, values are significantly different from 0, meaning that there is indeed a stronger attentional bias towards only one emotional face at a time, when both were probed in the same trial (p = 0.019).

4. Discussion

4.1 VWM biasing effect on visual attention for multiple items is here, yet not time specific

The way what we are temporarily storing in our Visual Working Memory can influence and drive our everyday visual search is surely a deep and broad question. We chose to pursue it by analyzing microsaccades with an eye-tracker, and by adopting an RSVP visual search task similar to the one used by Olmos-Solis et al. (2017), but with multiple search templates to be formed at the same time.

We questioned if multiple VWM search templates can bias visual attention, expecting to find leftward microsaccades amplitude when the memorized items were matching to the *left*, and rightward microsaccades when the items were matching to the *right*. The results we retrieved seem to show that multiple VWM content *does* bias microsaccades when the memorized items were matching to the left or to the right during RSVP presentation, even if participants were instructed to keep their gaze on the central fixation cross (*see* Figure 1). However, the interaction between the left and right matching probes with the 4 RSVP time conditions did not turn to be significant. That means that two VWM search templates *do* bias visual attention (contrasting with the findings in Van Moorseelar et al., 2014), but the biasing effect we observed is not time-specific (as opposed to what has been found by Olmos-Solis et al., 2017). Thus, we proceeded to reject our first null hypothesis (Experiment 1, H_1). Future research may deeper explore how multiple VWM content biases attention over time; for instance, the time grouping conditions (i.e., early, middle, and late) could be revisited for deeper insights yet to come.

The patterns of the average microsaccades amplitude when items are matching to the left, and to the right (*see* Figure 6), seem to mirror the already mentioned findings about how microsaccades can track general and search template-specific preparation for visual search (Olmos-Solis et al., 2017). Specifically, subtle eye movements can track attentional allocation prior to the visual search task to appear with task-relevant representation in VWM, even if subjects have been instructed to maintain their gaze into the fixation point (Van Loon et al., 2017); however, the evidence we retrieved on the matter was not significant, as we did not find any biasing effects in the interaction between left and right matching probes with early, middle, and late RSVP time conditions.

We questioned if there was any difference between Search Template A or B to bias more attention, whenever both were probed in the same trial. In other words, if the two VWM templates were biasing attention in the same way or if only one of them (either A or B) was biasing more microsaccades. We found that both items A and B were biasing attention in a similar way. Therefore, we did not reject our second null hypothesis (Experiment 1, H_2).

Lastly, by looking at the general performance for all participants, the widest difference between the proportion of correct answers was between yellow and lilac (as A: 14.3%; as B: 10.6%).

Just as importantly, we acknowledged a potential design limitation for Experiment 1: the visual search task answer time (1s) or item presentation (0.1s) were probably too short. A total of seven participants had been discarded because of poor performance or missing answers to the visual search task. That may also reflect the lower overall performance if compared to Experiment 2.

4.2 VWM attentional bias for emotional faces is stronger for only one search template

Colored cues are surely an essential part of our everyday lives when activating visual search. They may be part of any tool design, user interface element, advertising, and so on. However, we also experience the need to elaborate and process stimuli that are way more complex than colors alone.

That is the reason why we ran a second experiment with emotional and neutral human face stimuli to dive deeper into investigating the possibility of the biasing effect depending on the nature of the stimuli itself (i.e., simple vs. complex), and to eventually compare results with those from Experiment 1. Here we found an effect on the interaction between left and right matching probes and RSVP timewindows, which was not detected with the post-hoc test we ran afterwards. We did not find a biasing effect within left and right matching faces conditions, nor in their relation to time, making us refusing to reject our first null hypothesis for Experiment 2 (H_1).

Given the previously discussed findings about memory facilitations for certain emotional faces, we also expected to find differences in task performance for different emotions and/or neutrality within participants. We thought to observe generally wider differences in the proportion of correct answers given, with higher scores for fearful and angry faces – as it has been shown for them to be easier to retain in VWM (e.g., Gambarota & Sessa, 2019; Jackson et al., 2013; Lee & Cho, 2019; Lingxia et al., 2016). However, we did not find large differences in search task performance between different facial emotions and/or neutrality (*see* Figure 8); moreover, some participants (*N*: 5) reported that fearful faces were the hardest to memorize. The widest difference we found in performance was between happy and fearful faces, of 7.5% when probed as A and 1.4% as B (in terms of proportion of correct answers).

The patterns from the overall averaged microsaccade amplitude during RSVP we observed in Experiment 2 seemed to suggest the presence of the well-known pseudo-neglect effect, for which healthy subjects slightly and systematically shift their spatial attention to the left (e.g., Bowers & Heilman, 1980; Jewell & McCourt, 2000). Indeed, there is a quite consistent bias to the left throughout all the RSVP, with its highest peak between the 12^{th} and 17^{th} RSVP timewindows (*see* Figure 9), with more negative values to account for a smaller amplitude strength to the right (i.e., participants making microsaccades to the left even if the memorized faces were matching to the right).

However, it may be that these observations could not fully explain the patterns we found; that may also be due to the fact that, as reported in Jewell & McCourt's meta-analysis (2000), it is less likely to observe pseudo-neglect effects in relatively young, female subjects (McCourt et al., 1997). In fact, ⅔ of our participant pool for Experiment 2 consisted of females aged between 19 and 26. Even if the sample characteristics are similar to the ones we analyzed in Experiment 1, we did not find any trace of a pseudo-neglect effect in our previous experiment. Future research investigating the presence or absence of pseudo-neglect effects in visual search with multiple items would surely be interesting to expect.

When moving to our second hypothesis, we followed the same logical steps we took in Experiment 1 for data analysis and significance testing. We found a significant effect for only one Search Template to bias more microsaccades. A possible explanation to this could lie in the very nature of the stimuli we presented in Experiment 2: according to Caharel et al. (2014), ~200ms would be needed to tell whether a face is familiar or not; that is because there are several facial features to process holistically. Thus, in a context of fast visual search for multiple human emotional and neutral faces (i.e., 100ms) with limited attentional resources, there might be a *priority* being set for only one search template to bias attention. Simply put, participants could not be biased by two emotional faces at the same time, but just one. Given the empirical data we retrieved, we rejected our second null hypothesis for Experiment 2 (H_2).

Lastly, we would like to point out once again that the sample size for Experiment 2 is smaller than the one we analyzed for Experiment 1, and that the design of the experiment itself has been slightly changed to fit more complex stimuli (i.e., human emotional and neutral faces) within the same paradigm. Therefore, any reflection on the discussed similarities and differences with Experiment 1 should also take into account these design discrepancies.

5. Conclusions

Ultimately, we may conclude that VWM content biases visual attention when handling two search templates at the same time (in contrast with findings by Van Moorseelar et al., 2014); contrary to Olmos-Solis et al. (2017) findings, the biasing effect we observed was not time-specific. Furthermore, we also found that there is an attentional bias preference (or priority) being set for just one item when two human emotional faces are stored by VWM at the same time. Future research deeper investigating details about the relationship between VWM attentional bias with multiple items and time would be interesting to expect.

Acknowledgements

We would like to thank the Department of Social and Behavioral Sciences of Utrecht University for the research founding, including participants compensations and the equipment for data collection (excluding the free meal "Pasta alla Carbonara" contest for the best performer). We would also like to mention the "Empirical Foundations for Eye Tracking Reporting" guideline by Holmqvist et al., 2021, which we took under advise during data collection and reporting, hopefully for better comparability and replicability with other and future eye tracking research on the matter.

Supplementary Material

Pre-experiment training presentations

As we stated earlier in this paper, the pre-experiment training consisted in a 2-minute PowerPoint presentation to explain the trial structure to the participants, before starting 10 practice trials and the main experiment.

Hereby the link to the training .PPT for Experiment 1:

- Thesis – Experiment 1 – Tutorial.

And here is the link for the second experiment presentation:

- Thesis – Experiment 2 – Tutorial.

References

[1] AWH, E., & JONIDES, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, **5** (3), 119 –126.

[2] BOWERS D, HEILMAN KM (1980). Pseudoneglect: effects of hemispace on a tactile line bisection task. *Neuropsychologia*, **18**, 491-498.

[3] CAHAREL, S., RAMON, M. & ROSSIN, B. (2014). Face familiarity decisions take 200ms in the human brain: Electrophysiological evidence from a go/non-go speeded task. *Journal of Cognitive Neuroscience*, **26**, 81-95.

[4] CALVO, M., & NUMMENMAA, L., & AVERO, P. (2008). Visual Search of Emotional Faces. *Experimental Psychology* (Zeitschrift für Experimentelle Psychologie), **55**. 359-370. doi: 10.1027/1618-3169.55.6.359.

[5] DESIMONE, R., & DUNCAN, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, **18**, 193–222.

[6] DOWNING, P. E. (2000). Interactions between visual working memory and selective attention. *Psychological Science*, **11** (6), 467– 473.

[7] EASTWOOD, J., SMILEK, D., & MERIKLE, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception and Psychophysics*, **64**, 1004–1013.

[8] ENGBERT, R. (2006). Microsaccades: A microcosm for research on oculomotor control, attention, and visual perception. *Progress in Brain Research*, **154**, 177–192.

[9] EYSENCK, M. W., KEANE M. T. (2020). Face Recognition. *Cognitive Psychology. A student's handbook*, Routledge, Psychology Press, New York, **8**th ed., 116-130.

[10] FOLK, C., REMINGTON, R. W., & JOHNSTON, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 1030 – 1044.

[11] FOLK, C. L., LEBER, A. B., & EGETH, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Perception & Psychophysics*, **64** (5), 741–753.

[12] FOX, E., LESTER, V., RUSSO, R., BOWLES, R., PICHLER, A., & DUTTON, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition and Emotion*, **14**, 61–92.

[13] GAMBAROTA F, SESSA P. (2019) Visual Working Memory for Faces and Facial Expressions as a Useful "Tool" for Understanding Social and Affective Cognition. *Frontiers of Psychology*, **10**, 2392. doi: 10.3389/fpsyg.2019.02392.

[14] GAZZANIGA, M. S., IVRY, B. R., MANGUN, G. R. (2019a). Mechanisms of Memory. *Cognitive Neuroscience. The biology of the mind*, International Student Edition, W. W. Northon & Company, London, **5**th ed., 384-395.

[15] GAZZANIGA, M. S., IVRY, B. R., MANGUN, G. R. (2019b). Selective Attention and the Anatomy of Attention. *Cognitive Neuroscience. The biology of the mind*, International Student Edition, W. W. Northon & Company, London, **5**th ed., 276-278.

[16] HOLLINGWORTH, A., MATSUKURA, M., & LUCK, S. J. (2013). Visual Working Memory Modulates Rapid Eye Movements to Simple Onset Targets. *Psychological Science*, **24** (5), 790–796. doi: 10.1177/0956797612459767.

[17] HOLMQVIST, K., ÖRBOM, S., HOOGE, I., NIEHORSTER, D., ALEXANDER, R., ANDERSSON, R., BENJAMINS, J., BLIGNAUT, P., BROUWER, A.-M., CHUANG, L., DALRYMPLE, K., DRIEGHE, D., DUNN, M., ETTINGER, U., FIEDLER, S., FOULSHAM, T., GEEST, J., HANSEN, D., HUTTON, S., HESSELS, R. (2021). Eye tracking: empirical foundations for a minimal reporting guideline. *Behavior Research Methods*. doi: 10.3758/s13428-021-01762-8.

[18] JACKSON, M., & LINDEN, D., & RAYMOND, J. (2013). Angry expressions strengthen the encoding and maintenance of face identity representations in visual working memory. *Cognition & emotion*, **28**. doi: 10.1080/02699931.2013.816655.

[19] THE JAMOVI PROJECT (2022). *jamovi*. (Version 2.3) [Computer Software]. Retrieved from https://www.jamovi.org.

[20] JEWELL G. & MCCOURT M. E. (2000) Pseudoneglect: a review and meta-analysis of performance factors in line bisection tasks, *Neuropsychologia*, **38** (1), 93-110. doi: 10.1016/S0028-3932(99)00045-7.

[21] LAUBROCK, J., ENGBERT, R., KLIEGL, R. (2005). Microsaccade dynamics during covert attention. *Vision Research*, **45** (6), 721–730.

[22] LEE, H.J., CHO, Y.S. (2019) Memory facilitation for emotional faces: Visual working memory trade-offs resulting from attentional preference for emotional facial expressions. *Memory & Cognition,* **47**, 1231–1243. https://doi.org/10.3758/s13421-019-00930-8.

[23] LENTH, R. (2020). *emmeans:* Estimated Marginal Means, aka Least-Squares Means. [R package]. Retrieved from https://cran.r-project.org/package=emmeans.

[24] LINGXIA, F., CODY, D., RENLU, G., MENGSI, X., LIUTING, D., DONG, Y. (2016) Visual working memory representations guide the detection of emotional faces: An ERP study. *Vision Research*, **119**, 1-8. doi: 10.1016/j.visres.2015.08.021

[25] LUNDQVIST, D., FLYKT, A., & ÖHMAN, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

[26] MARTINEZ-CONDE, S., OTERO-MILLAN, J. & MACKNIK, S. (2013). The impact of microsaccades on vision: towards a unified theory of saccadic function. *Nature Reviews Neuroscience*, **14**, 83–96. doi: 10.1038/nrn3405.

[27] MCCOURT M.E., MARK, V.W., RADONOVICH K.J., WILLISON, S.K., FREEMAN P. (1997). The effects of gender, menstrual phase and practice on the perceived location of the midsagittal plane. *Neuropsychologia*, **35**, 717-724.

[28] MOORES, E., LAITI, L. & CHELAZZI, L. (2003). Associative knowledge controls deployment of visual selective attention. *Nature Neuroscience*, **6**, 182–189.

[29] ÖHMAN, A., LUNDQVIST, D., & ESTEVES, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, **80**, 381–396.

[30] OLIVERS, C. N. (2009). What drives memory-driven attentional capture? The effects of memory type, display type, and search type. *Journal of Experimental Psychology: Human Perception and Performance*, **35** (5), 1275-1291. doi: 10.1037/a0013896.

[31] OLIVERS, C. N., PETERS, J., HOUTKAMP, R., ROELFSEMA, P. R. (2011). Different states in visual working memory: when it guides attention and when it does not. *Trends in Cognitive Science*, **15** (7), 327-334. doi: 10.1016/j.tics.2011.05.004.

[32] Olmos-Solis, K., Van Loon, A. M., Los, S. A., & Olivers, C. N. L. (2017). Oculomotor measures reveal the temporal dynamics of preparing for search. *Progress in Brain Research*, **236** (1), 23. doi: 10.1016/bs.pbr.2017.07.003.

[33] Pastukhov, A., Braun, J. (2010). Rare but precious: microsaccades are highly informative about attentional allocation. *Vision Research*, **50** (12), 1173–1184.

[34] Pinkham, M. G., Baron R., Sasson N. J., R. Gur C. (2010). The face in the crowd effect: Anger superiority when using real faces and multiple identities. *Emotion*, **10** (1), 141-146.

[35] Phelps, E. A. (2004). Human emotion and memory: interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology*, **14** (2), 198-202.

[36] R Core Team (2021). *R*: A Language and environment for statistical computing. (Version 4.1) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from MRAN snapshot 2022-01-01).

[37] Schreij, D., Los, S. A., Theeuwes, J., Enns, J. T., & Olivers, C. N. L. (2014). The interaction between stimulus-driven and goal-driven orienting as revealed by eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, **40** (1), 378–390.

[38] Sessa, P., Luria, R., Gotler, A., Jolicoeur, P., & Dell'Acqua, R. (2011). Interhemispheric ERP asymmetries over inferior parietal cortex reveal differential visual working memory maintenance for fearful versus neutral facial identities. *Psychophysiology*, **48**, 187-197.

[39] Singmann, H. (2018). *afex*: Analysis of Factorial Experiments. [R package]. Retrieved from https://cran.r-project.org/package=afex.

[40] Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, **6**, 11-20.

[41] Soto, D., & Humphreys, G. W. (2008). Stressing the mind: The effect of cognitive load and articulatory suppression on attentional guidance from working memory. *Perception & Psychophysics*, **70**, 924–934. doi: 10.3758/PP.70.5.924.

[42] Sreenivasan, K. K., & Jha, A. P. (2007). Selective attention supports working memory maintenance by modulating perceptual processing of distractors. *Journal of Cognitive Neuroscience*, **19**, 32– 41.

[43] Van Loon, A.M., Olmos-Solis, K., Olivers, C.N.L. (2017). Subtle eye movement metrics reveal task-relevant representations prior to visual search. *Journal of Vision*, **17** (6), 13.

[44] Van Moorselaar, D., Theeuwes, J., & Olivers, C. N. L. (2014). In Competition for the Attentional Template: Can Multiple Items Within Visual Working Memory Guide Attention?. *Journal of Experimental Psychology: Human Perception and Performance*, advance online publication. doi: 10.1037/a0036229.

[45] Watson, T. L., Robbins, R. A. (2014). The nature of holistic processing in face and object recognition: Current opinions. *Frontiers in Psychology*, **5** (3).

[46] Williams, M. A., Moss, S. A., Bradshaw, J. L., & Mattingley, J. B. (2005). Look at me, I'm smiling: Visual search for threatening and nonthreatening facial expressions. *Visual Cognition*, **12**, 29–50.

[47] Zhang, B., Zhang, J. X., Huang, S., Kong, L., & Wang, S. (2011). Effects of load on the guidance of visual attention from working memory. *Vision Research*, **51**, 2356–2361.