ANGRY EXPRESSIONS STRENGTHEN THE ENCODING AND

MAINTENANCE OF FACE IDENTITY REPRESENTATIONS IN VISUAL

WORKING MEMORY

Author's PREPRINT COPY

Margaret C. Jackson[1], David E. J. Linden[2] and Jane E. Raymond[3]

[1]*University of Aberdeen, UK*
[2]*Cardiff University, UK*
[3]*University of Birmingham, UK*

Correspondence to:

Dr Margaret C. Jackson
School of Psychology
William Guild Building
University of Aberdeen
Aberdeen
AB24 3FX
UK
Email: m.jackson@abdn.ac.uk
Tel: +44 (0)1224 272236
Fax: +44 (0)1224 273426

Short title: Angry faces in working memory

Key words: Faces, Anger, Emotion, Working Memory, Maintenance, Encoding

Word count = 9144

Acknowledgements

ABSTRACT

Visual working memory (WM) for face identities is enhanced when faces express negative versus positive emotion. To determine the stage at which emotion exerts its influence on memory for person information, we isolated expression (angry / happy) to the encoding phase (Experiment 1; neutral test faces) or retrieval phase (Experiment 2; neutral study faces). WM was only enhanced by anger when expression was present at encoding, suggesting that retrieval mechanisms are not influenced by emotional expression. To examine whether emotional information is discarded on completion of encoding or sustained in WM, in Experiment 3 an emotional word categorization task was inserted into the maintenance interval. Emotional congruence between word and face supported memory for angry but not for happy faces, suggesting that negative emotional information is preferentially sustained during WM maintenance. Our findings demonstrate that negative expressions exert sustained and beneficial effects on WM for faces that extend beyond encoding.

The human face conveys both person identity and social-emotional information within the same stimulus. Static, structural information allows us to individuate others so that, using long-term memory, we can link prior experience with current events concerning that particular individual. In contrast, transient movements of facial muscles, leading to facial expressions, communicate temporary emotional information about that person's internal state and provide important clues as to his or her immediate and future social intent. The fleeting nature of emotional expressions coupled with the static nature of face identity information means that within a typical social episode, emotional expression and identity information are rarely fully contemporaneous. Yet, keeping track of the relevant players and their emotional states over time during a social interaction is critical for planning appropriate behavioural responses. This suggests that short-term or working memory (WM) for emotional expression information may play an important role in the visual social cognition of human faces.

Several studies have shown that negative emotional expression facilitates visual WM for face identity (Jackson, Wolf, Johnston, Raymond, & Linden, 2008; Jackson, Wu, Linden, & Raymond, 2009; Sessa, Luria, Gotler, Jolicoeur, & Dell'Acqua, 2011). Such studies measured WM performance using a simple delayed discrimination task in which a study array comprising a small number of different people's faces, all bearing the same expression, is presented for several seconds for encoding. After a one second maintenance interval, a face from the study array (match-trial condition) or a different person's face (non-match-trial condition) is presented. The task is to report whether the identity of the test face is the same as or different from one seen in the preceding study array. The primary finding is that WM for the identity of faces bearing angry (Jackson et

al., 2008; 2009) or fearful (Sessa et al., 2011) facial expressions is significantly better than WM for faces bearing happy or neutral expressions. WM for happy faces was not significantly different from that for neutral faces (Jackson et al., 2008, 2009), indicating that an angry expression specifically boosts WM whereas as a happy expression has no measurable effect. Related research has shown that anger-specific enhancement of WM may be related to dopaminergic processes. It is correlated with increased activity in the basal ganglia, specifically the globus pallidus (Jackson et al., 2008), and requires an increase in dopamine levels to become observable in patients with Parkinson's Disease (Subramanian, Hindle, Jackson, & Linden, 2010). The effect of face expression on WM for face identity is consistent with a growing body of evidence showing that mechanisms supporting face identification interact with those supporting expression interpretation (Gallegos & Tranel, 2005; Galster, Kahana, Wilson, & Sekuler, 2009; Ganel & Goshen-Gottstein, 2004; Ganel, Valyear, Goshen-Gottstein, & Goodale, 2005; Kaufmann & Schweinberger, 2004; Vuilleumier, Armony, Driver, & Dolan, 2001; see Vuilleumier & Pourtois, 2007, for a review) and is contrary to a widely accepted and traditional face perception model (Bruce & Young, 1986) that posits independence of these systems.

Although anger-specific effects on WM are robust and replicable, it remains unclear how memory and emotional expression perception processes interact to yield this effect. The goal of the current series of experiments was to further investigate emotion-specific enhancement of visual WM by determining at which stage or stages of the WM process emotional expression influences memory. Such studies have the potential to advance understanding of how emotional information modulates high-level visual processes, especially those supporting fluent social-emotional cognition.

       Visual WM is traditionally viewed as a flexible but limited capacity memory mechanism that operates over a few seconds to enable mental access to visual representations of stimuli after they are no longer visible (Cowan, 2000; Luck & Vogel, 1997). It is widely assumed to comprise three stages of processing. First, to-be-remembered items must be perceptually encoded into a mental representation. Experimentally, this occurs during the study phase when an array of items is presented for scrutiny. Second, when the to-be-remembered items are no longer present (a period referred to as the maintenance interval), their representations must be maintained in WM so they can be accessed later. Third, stored representations must be retrieved and compared with any newly available sensory information (test stimulus) to recognize the reappearance of just previously seen items or to determine if the visual scene has changed. The brain makes use of visual WM processes every time the eyes or head move, or when an object in the external word undergoes movement or transformation (e.g., when the expression on a face changes), thus enabling scenes to appear coherent and stable in the face of retinal image changes (Henderson & Hollingworth, 2003).

       We conducted three different experiments using the delayed face discrimination task of previous studies to test the impact of expression on WM performance during the different WM phases. In previous studies the test face always had the same expression as faces seen during the study phase and, on match-trials, exactly matched one of the study array face images. Here, to better identify at which stage of the WM process expression exerts its influence, we modified the paradigm so that expression was present during encoding (angry versus happy study array) but not at retrieval (neutral test face) (Experiment 1). In Experiment 2, the faces were neutral during encoding but the test face

was expressive (angry or happy). A difference in accuracy on this task when angry versus happy faces were presented was used to index emotion-specific enhancement effects on WM performance.

The goal of these first two experiments was to determine whether emotion-specific WM enhancement depended on the presence of a negative emotional facial expression during encoding or during retrieval. The rapid capture of attention by threatening versus non-threatening stimuli is widely documented in the attention literature (Bannerman, Milders, & Sahraie, 2010; Eastwood, Smilek, & Merikle, 2003; Feldmann-Wustefeld, Schmidt-Daffy,& Schubö, 2011; Fenske & Eastwood, 2003; Fox & Damjanovic, 2006; Hahn, Carlson, Singer & Gronlund, 2006, Horstmann, Borgstedt & Heumann, 2006; Huang, Chang & Chen, 2011), particularly in relation to angry faces. This suggests that, within the context of the current task, a face portraying a threatening expression at retrieval might receive enhanced attention that could in turn enhance its perceptual processing (Kastner & Ungerleider, 2000; Carrasco, Ling & Read, 2004), thereby facilitating the comparison processes underlying retrieval. Thus, a candidate mechanism that could contribute to emotion-specific WM enhancement is an attention boost occurring during presentation of the test face. This possibility was examined in Experiment 1 by presenting emotionally expressive faces (angry versus happy) in the study array followed by a neutrally expressive face at test. We did not test WM for neutral faces as it has been previously shown in numerous experiments that WM for faces with happy expressions is non-significantly different from that for faces with neutral expression (Jackson et al., 2008, 2009). If emotion-specific enhancement effects in WM depend solely on a boost to retrieval processes, then WM performance should be

unaffected by the expression in study array faces. Alternatively, expression effects on WM in this experiment would demonstrate that emotional expression during the encoding phase alone is sufficient to boost visual WM for face identity. In the second experiment these conditions were reversed such that only the test stimulus bore an emotional expression (angry or happy) and the study array faces were always neutral. If the presence of threat at retrieval was sufficient to boost recognition, then an emotion-specific enhancement effect on WM should be observed.

An important feature of these experiments is that the test and study faces always had different expressions. All previous studies of emotion-specific WM enhancement presented the same photo exemplars at study and test on match-trials. An essential characteristic of visual WM is its capacity to tolerate modest mismatch or transformation between initially encoded stimuli and subsequent viewed images. This is in contrast to a putative lower-level visual short-term memory (STM) cache that requires exact correspondence between encoded and subsequently seen stimuli to produce recognition (Atkinson & Shiffrin, 1968). Whereas WM would allow one to recognize a previously viewed smiling face as the same person currently frowning, STM would not. Therefore, this new procedure in which study and test faces always portray a different expression additionally allowed us to determine whether the emotion-specific enhancement effect on memory for face identity is mediated by WM or by a simpler STM mechanism.

The aim of Experiment 3 was to examine whether task-irrelevant emotional expression information is discarded once encoding is complete, or whether emotion is maintained in WM despite the absence of facial expression during the maintenance interval and at retrieval. Sessa et al. (2011) compared visual WM for fearful versus

neutral face identities, whilst measuring the sustained posterior contralateral negativity (SPCN) component of visual event-related potentials (ERPs) time locked to the encoding display. This component (also known as contralateral delay activity; CDA) is considered to index visual WM maintenance processes (Klaver, Talsma, Wijers, Heinze, & Mulder, 1999; McCollough, Machizawa, & Vogel, 2007; Robitaille, Grimault, & Jolicoeur, 2009; Vogel & Machizawa, 2004) and was enhanced when the study array comprised fearful versus neutral faces. These findings provide evidence that visual WM maintenance processes are indeed influenced by the presence of negative emotion at encoding, but do not directly inform us as to whether emotional information is maintained in WM in addition to face identity information.

The contents of WM during maintenance are shown to be fragile and susceptible to general interference. Additional information presented during the maintenance period of a WM task has been shown to impair memory accuracy. For example, De Fockert, Rees, Frith, and Lavie (2001) showed that presenting to-be-ignored distracter faces during a high load digit WM task impaired digit recall. There is also evidence that the valence of distracting information impacts on WM for neutral information. Negative (but not neutral) distracter stimuli have been shown to impair WM for neutral words (Buchner, Rothermund, Wentura, & Mehl, 2004) and for faces with a neutral expression (Dolcos, Iordan, & Dolcos, 2011; Dolcos & McCarthy, 2006). These findings suggest that incidental negative information detracts attention from an ongoing neutral task in which WM is engaged. It has also been shown that neutral intervening information that is task-relevant and which requires an immediate response can influence WM for other neutral stimuli. Yoon, Curtis, and D'Esposito (2006) found that WM for faces was

impaired when a face versus scene recognition task was inserted into the WM maintenance period, suggesting that intervening information that was perceptually or categorically similar to the contents of WM served to interfere with maintenance processes. Studies such as these imply that the strength of stimulus representations maintained in WM can be degraded in some way either by a general reduction in attention or by more specific perceptual interference mechanisms.

It is also possible that emotional information retained in WM can be influenced by intervening higher-level, conceptual information such as valence. Recent work provides evidence that the valence of task-irrelevant intervening information can influence WM for emotional stimuli. Jackson, Linden, and Raymond (2012) found that when the WM maintenance period was protracted to 9 seconds, the angry versus happy benefit observable with a 1 second maintenance period (Jackson et al., 2009) was abolished. However, when a task-irrelevant emotional (positive or negative) word was presented during maintenance, WM for the identity of angry faces was boosted (relative to when a neutral or no word was presented), resulting in the re-instatement of the anger-enhancement effect. WM for happy faces was unaffected by the presence or absence of differently valenced words.

In Experiment 3, we aimed to directly probe the emotional content of visual WM in our faces task (i.e., to ask whether emotional information is maintained or not) by presenting task-relevant information during the maintenance interval that was either similar or dissimilar in valence to the emotional expression conveyed by the faces at encoding. To do this, we used the same delayed discrimination task as in Experiment 1 (angry or happy faces at study, a neutral face at test), but additionally presented a positive

or negative word during a 2 second maintenance interval. Participants categorized the word as positive or negative as quickly and accurately as possible before making a retrieval response to the WM test face which appeared shortly after. Applying the categorization task to the words ensured that attention was allocated to the words and their valence, a factor that was not controlled in the Jackson et al. (2012) study.

Considering previous findings of perceptual interference effects in WM when intervening items are task-relevant (Yoon et al., 2006), we reasoned that if emotional information from the study array was maintained, then an emotionally congruent word presented during the maintenance interval (e.g., a negative word following a study array of angry faces) might interfere with WM for faces and an incongruent word should facilitate it. Furthermore, WM for faces might be specifically supported by the sustained activation of negative (angry) but not positive (happy) emotional information over time, in order that an effective and immediate response to threat can be prepared and executed. If negative facial emotional information is maintained more strongly than happy information, then any emotional congruence effect of words on WM performance should be greater when faces at study express anger than when they express happiness. Alternatively, it is possible that the presence of negative emotion boosts the initial encoding of stable person information, resulting in more precise representations of face identity that are thus better maintained over time without the need to maintain emotional expression information in WM. If facial emotion is not maintained in visual WM beyond the encoding phase, then word valence should have no effect on visual WM performance, regardless of the emotion expressed in the study array.

GENERAL METHODS

Participants

Adult volunteers recruited through the University student and community panels received course credits or money in exchange for participation. All reported normal or corrected-to-normal eyesight. APA ethical standards for treatment of human volunteers were upheld.

Apparatus

Stimuli were presented on a 22-inch Mitsubishi Diamond- Pro 2060u monitor (32-bit true color; resolution 1280 x 1024 pixels) using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) and were viewed from a distance of 50 cm.

Stimuli

Grayscale images of six different male individuals, each displaying an angry, happy, and neutral expression (18 faces stimuli in total) were selected from the Ekman and Friesen (1976) database. Faces subtended approximately 1.43^{o} x 1.36^{o} degrees of visual angle. Hair was removed by cropping each face into an oval in order to minimize reliance on featural cues to perform the task. Visual WM study arrays consisted of between one and four faces (depending on the Experiment) displayed in a 2 x 2 grid. The distance between faces was 0.38° visual angle (0.2 cm) on the horizontal and vertical axes and 1.67° (1.2 cm) on the diagonal axis. Face location within the grid was randomized. Scrambled grayscale faces filled study display locations when fewer than four faces were presented. Eight different scrambled images were created by segmenting eight different faces into 25 squares and then randomly rearranging them. The composite image was cropped into an oval pattern to maintain a face-like outline. All faces within

each study array bore the same expression, either angry or happy (Experiments 1 and 3), or neutral (Experiment 2). In Experiments 1 and 3 the test face was always neutral and in Experiment 2 the test face was either angry or happy. Previous Self-Assessment Manikin (Bradley & Lang, 1994) ratings of arousal for the same face images used here showed that angry faces were perceived as similarly arousing as happy faces (mean ratings were 0.94 ± 0.44 and 0.83 ± 0.46 for angry and happy faces respectively and the difference was non-significant; Jackson et al., 2009, pp 367).

A set of consonants was used for the verbal suppression task in Experiments 1 and 2. No verbal suppression was used in Experiment 3.

Design and Procedure

The main components of the trial sequence for each experiment are shown in Figure 1. Each trial began with a 3000 ms fixation cross that briefly grew in size to signal the start of a trial, followed by a 2000 ms presentation of a WM study array of 1, 2, 3, or 4 faces (angry or happy in Experiments 1 and 3, neutral in Experiment 2). Only two faces were presented for study in Experiment 3. When the study array terminated, a blank maintenance interval comprising only the central fixation cross was presented. This lasted 1000 ms in Experiments 1 and 2, but was extended in Experiment 3 (See Experiment 3 for details). Then a single test face was presented centrally until response. On half of trials, the test face shared identity with one of the faces presented in the preceding study array (match condition); on remaining trials it shared identity with none of the study faces (non-match condition). The task was to respond 'yes' if the test face identity was present in the preceding study array or 'no' if it was not. In Experiments 1 and 2, trials were split into 8 blocks (32 trials per block; 256 trials in total). There were 64 angry and 64 happy

trials, of which there were 16 trials in each study array size (WM load) and condition

(50% match). Emotion, WM load, and match/non-match conditions were presented in a

pseudo-random order.

       In Experiments 1 and 2 a verbal suppression task was administered concurrently

to minimize the use of verbal labels (see Jackson et al., 2009). At the start of each block

two letters were presented, with participants instructed to repeat the letters out loud

during the entire block. To check verbal rehearsal, participants stated whether two letters

presented at the end of each block were the same or different to the ones they had been

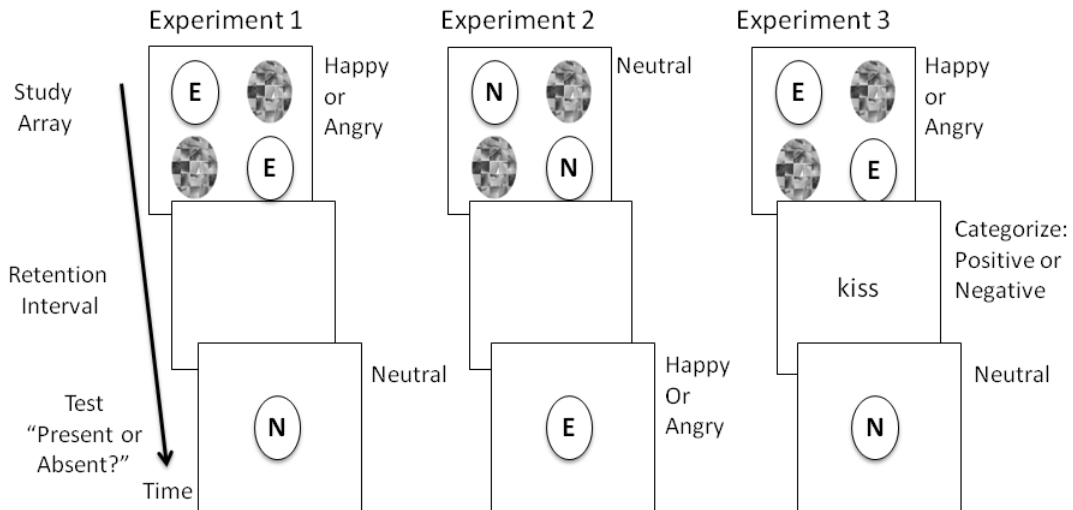repeating. Participants performed above 80% correct on the verbal task in all

experiments.

Figure 1. An illustration of an example trial for each experiment. Ovals with the letter E inside denote an emotional (happy or angry) face, and ovals with the letter N inside denote a neutral face. After a brief fixation display (not shown), a study array of faces was presented for 2000 ms followed by a maintenance interval (1000 ms for Experiment 1 &2; 2000 ms for Experiment 3). This interval was blank for Experiment 1 and 2, but in Experiment 3 it contained a 100 ms presentation of either a positive or negative word that the participant categorized (positive or negative) as quickly as possible. The word appeared either 250 ms or 1000 ms after study array offset. Finally, a test face appeared;

participants reported whether the face had been present or absent in the study array. All faces in the study array shared the same expression: Happy or angry in Experiments 1 and 3; neutral in Experiment 2. The test face was always neutral in Experiment 1 and 3, but either happy or angry in Experiment 2.

Pilot testing showed that this task was significantly more difficult than the original study in which all faces showed the same emotion at encoding and retrieval. To facilitate performance, participants were allowed at the start of the session to study the faces used in the experiment for an unlimited amount of time, examining how each individual appeared with an angry, happy, and neutral expression. The study phase is unlikely to have facilitated WM by the creation of long-term memory traces for the faces (see Jackson & Raymond, 2008) because performance on this task still remained markedly lower (55% lower on average) than the original experiments in which no study phase was allowed. Then, 20 practice trials were given: 10 practice trials with an encoding period of 4000 ms (to aid task learning), and then another 10 using a study interval matching that used in the main experiment.

Data Analysis

Sensitivity for recognizing the test face identity was calculated for each participant for each condition using d'. This was computed for each condition by subtracting the Z-transform of the proportion of False Alarms (FA; non-match trials on which participants incorrectly responded 'yes') from the Z-transform of the proportion of Hits (match trials for which participants correctly responded 'yes'). Analyses of variance (ANOVAs) were conducted on d' scores, hits, and false alarm (FA) rates using emotion (angry, happy) and - in Experiments 1 and 2 - load (1, 2, 3, 4) as within subject factors. ANOVAs conducted in Experiment 3 replaced load with other within-subject factors, as described there. Reaction times (RTs) to make the WM face retrieval response were also

analysed on correct trials only, trimmed to remove responses faster than 200 ms and slower than 4 standard deviations above the group mean in each Experiment. RT exclusions amounted to 0.50%, 1.08%, and 3.17% of trials in Experiments 1, 2, and 3 respectively. Alpha levels were set to .05 throughout.

EXPERIMENT 1

The aim was to test whether the presence of an angry expression in the test stimulus was necessary to produce an emotion-specific enhancement effect of WM. Additionally we sought to determine if this effect was dependent on the presentation of the same exemplar at test (on match-trials) or could withstand the use of a different test image of the same person, thereby showing this to be a WM rather than STM effect. Faces comprising each study array were either all angry or all happy and the test face was always neutral in expression.

Participants.

Twenty-two healthy individuals (17 females; mean age 19 years) took part.

Results.

As can be seen in Figure 2, performance indicated by d' values was significantly better when faces were angry versus happy during encoding ($F(1, 21) = 6.47$, $p = .02$). As expected, performance decreased significantly as the number of faces seen at study increased ($F(3, 63) = 37.38$, $p < .001$), but this factor did not interact with the effect of emotion ($F < 1$). On examination of hits and FA rates (see Table 1), we found that whereas hit rates were significantly modulated by emotion ($F(1, 21) = 9.94$, $p = .005$), FA rates were not ($F < 1$). The interaction between emotion and load was non-significant for both hit and FA rates ($F < 1$ in both cases). This pattern of results indicates that the

presence of an angry but not happy expression at encoding served to facilitate match decisions but emotion had a negligible effect on non-match decisions. Our hit and FA data are important because they allow us to ascertain that enhanced WM for angry faces is not simply due to greater perceptual similarity between an angry and a neutral face than between a happy and a neutral face. Previous work shows that happy faces are perceptually more discriminable from neutral faces than are angry faces (Mermillod, Vermeulen, Lundqvist, & Niedenthal, 2009), perhaps due to neutral faces appearing more negative (Lee, Kang, Park, Kim, & An, 2008). If our effects are driven by perceptual similarity rather than WM processes, then we should have seen significantly more false alarms in the angry versus happy face condition, but we did not.

An ANOVA on RTs (see Table 1) with emotion and load as within factors revealed a non-significant main effect of emotion ($F < 1$). There was a significant main effect of load ($F(3, 63) = 8.52$, $p < .001$) which reflects slowed RTs as load increased. The interaction between emotion and load was non-significant ($F(3, 63) = 1.91$, $p = .14$).
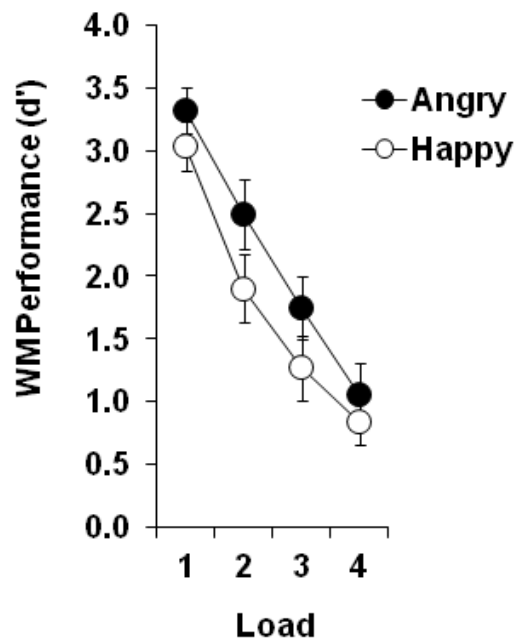
Figure 2. Group mean WM performance (d') scores obtained in Experiment 1 plotted as a function of the number of faces (load) in the study array. All study array faces had either angry or happy expressions in different conditions; the test face was neutral in expression. Vertical bars indicate ± 1 within-subject standard error.

These results replicate previous findings and additionally demonstrate that in the absence of emotional expression at retrieval, anger at encoding is sufficient to enhance visual WM. Furthermore, replication of the anger-enhancement effect when the face at retrieval was never a direct copy of a face at encoding, also serves to clarify that the original results were not simply due to a low level perceptual advantage in image matching afforded by angry faces, and demonstrates that the effect is mediated by WM rather than a lower-level visual STM mechanism that requires exact correspondence between study and test items to produce recognition.

A comparison between performance levels found here with those reported by Jackson et al. (2009, Experiment 1) shows that WM performance in the original study is better by around 1.0 d' unit when both study and test faces are expressive, versus when emotion is restricted to encoding. A mixed design ANOVA with study as a between factor and emotion and load as within factors, showed a significant main effect of study ($F(1, 44) = 21.30$, $p < .001$), but the two- and three-way interactions between study, emotion, and load were non-significant (all F's < 1). Clearly the task was significantly harder in the current experiment, likely due to the fact that participants were forced to actively extract face identity from emotional expression in order to successfully perform the task. However, this analysis shows that while removing emotion from the test face makes the overall task significantly harder, it does not impact on the angry face benefit. The magnitude of the anger (versus happy) enhancement effect is comparable in both this and the original experiment (0.40 and 0.48 d' units respectively). It is also important to note that the original study (Jackson et al., 2009) showed that WM for happy faces did not significantly differ from WM for neutral faces, suggesting that the presence of a positive expression neither enhances nor impairs WM relative to a neutral baseline. Thus, in this and previous studies, we interpret the finding of higher WM performance scores for angry than happy faces to reflect an anger enhancement rather than happy impairment effect.

Table 1. Mean Hits, FA rates, and RTs (ms) for the WM faces task in Experiments 1 and 2 as a function of face emotion and WM load. Standard errors are provided in brackets.

		Angry				Happy			
		L1	**L2**	**L3**	**L4**	**L1**	**L2**	**L3**	**L4**
Hits	Expt 1	0.94	0.84	0.77	0.71	0.88	0.74	0.70	0.63
		(0.01)	*(0.03)*	*(0.05)*	*(0.05)*	*(0.03)*	*(0.05)*	*(0.05)*	*(0.04)*
	Expt 2	0.91	0.72	0.69	0.63	0.81	0.71	0.64	0.60
		(0.04)	*(0.07)*	*(0.06)*	*(0.04)*	*(0.06)*	*(0.04)*	*(0.04)*	*(0.07)*
FAs	Expt 1	0.11	0.19	0.27	0.42	0.09	0.21	0.31	0.36
		(0.03)	*(0.04)*	*(0.04)*	*(0.05)*	*(0.02)*	*(0.04)*	*(0.04)*	*(0.05)*
	Expt 2	0.20	0.23	0.38	0.39	0.11	0.21	0.26	0.34
		(0.06)	*(0.06)*	*(0.05)*	*(0.04)*	*(0.04)*	*(0.05)*	*(0.03)*	*(0.07)*
RTs	Expt 1	1147	1222	1319	1243	1131	1192	1290	1311
		(66)	*(66)*	*(76)*	*(83)*	*(64)*	*(58)*	*(69)*	*(88)*
	Expt 2	1051	1207	1224	1314	1036	1135	1163	1252
		(77)	*(83)*	*(86)*	*(107)*	*(75)*	*(73)*	*(73)*	*(90)*

EXPERIMENT 2

Although the presence of threat at test may not be necessary for an emotion-specific enhancement effect on visual WM, an angry test face could nevertheless additionally enhance visual WM performance by eliciting an extra boost of attention that could facilitate the comparison process between the stored representations and the test image. To test this possibility, we used the same delayed discrimination task as in Experiment 1, except that here the study faces were always neutral and the test face was either happy or angry. The magnitude of change in the featural aspects of the image between study array and test presentation on same-trials was equivalent to that for Experiment 1.

Participants

Thirteen healthy individuals (8 females; mean age 28 years) took part. A power calculation using the angry versus happy effect size from Experiment 1 indicated that this sample size was sufficient.

Results and Discussion

As can be seen in Figure 3, WM performance was unaffected by the emotional expression of the test face ($F < 1$). Although performance decreased significantly as the number of faces seen at study increased ($F(3, 36) = 24.59$, $p < .001$), this factor did not significantly interact with test face emotion ($F < 1$). Furthermore, a mixed design ANOVA with Experiment (1 versus 2) as a between factor and emotion and load as within factors revealed a marginally significant interaction between emotion and experiment, $F(1, 33) = 3.63$, $p = .07$, indicating that removing expression from the study faces in Experiment 2 effectively diminished the effect seen in Experiment 1. Further analyses showed a non-significant effect of emotion on hit rates ($F(1, 12) = 1.77$, $p = .21$) and on FA rates ($F(1, 12) = 2.81$, $p = .12$), and the interactions between emotion and load for hits and FAs were non-significant (both F's < 1) (see Table 1). The lack of effect of emotional expression on FA rates in this experiment (and in Experiment 1) suggests that enhanced WM for angry faces found in Experiment 1 is not simply due to an 'emotion-induced recognition bias', an effect where negative items elicit a greater number of false positive reports (Johansson, Mecklinger, & Treese, 2004; Windmann & Kutas, 2001).
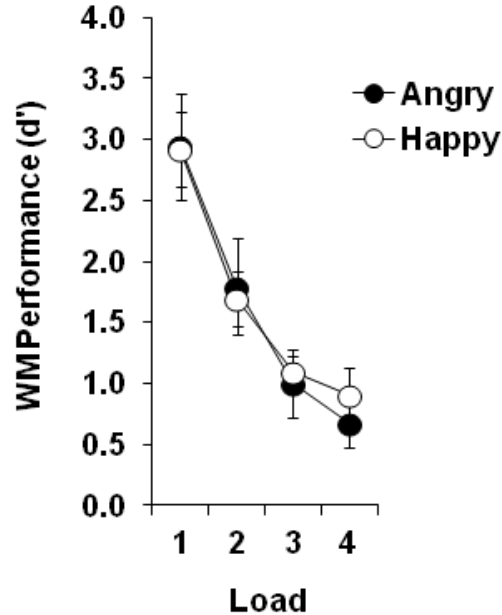
Figure 3. Group mean WM performance (d') scores obtained in Experiment 2 plotted as a function of the number of faces (load) in the study array. All study array faces were neutral in expression and, in different conditions, the test face was either angry or happy in expression. Vertical bars indicate ± 1 within-subject standard error.

An ANOVA on RTs (Table 1) with expression and load as within factors revealed a non-significant main effect of emotion ($F(1, 12) = 2.72$, $p = .13$). There was a significant main effect of load ($F(3, 36) = 10.87$, $p < .001$) which reflects slowed RTs as load increased. The interaction between emotion and load was non-significant ($F < 1$).

These results indicate that the presence of threat during the test phase does not appear to provide any benefit to performance on this task, and further suggest that in previous demonstrations of emotion-specific WM enhancement an emotional expression on the test face did not contribute to the effect. In conjunction with Experiment 1, these findings clearly indicate that emotion-specific enhancement effects on WM do not arise from processes related to retrieval and must be initiated during encoding.

EXPERIMENT 3

The question addressed in this experiment is whether the irrelevant emotional information available in the study array is maintained throughout the WM maintenance interval or is discarded once the study faces disappear. Participants performed the face WM task, as in Experiment 1, viewing expressive study array faces (load 2 only) and then a neutral test face. However, this time participants were required to categorize a word that was briefly presented during the WM maintenance interval as 'Positive' or 'Negative'. Of interest was the effect on visual WM performance of emotional congruence between words and face expression. Inserting the word task required lengthening the maintenance interval from 1 s to 2 s, in order to provide participants time to respond to the word. Additionally it afforded us an opportunity to manipulate word onset time thereby allowing us to probe whether retained emotional content decays within this interval.

Participants.

Twenty-five healthy individuals (17 females; mean age 24 years) took part. Data from three participants who had high error rates (> 10%) on the word task were excluded, leaving a total sample of 22.

Stimuli

Ten Positive-Negative word pairs were chosen for the word valence-categorization task: Smile-Frown; Kiss-Kick; Love-Hate; Agree-Argue; Laugh-Shout. Word length and frequency values (obtained from the Corpus of Contemporary American English) for each word within a pair were matched and each word in each pair was a conceptual opposite of the other. The word pairs were specifically chosen to reflect valence that ranged from a reference to facial expression (Smile-Frown), to physical acts

of emotion (Kiss-Kick, Laugh-Shout), to broader emotional concepts (Love-Hate, Agree-Argue) in order to avoid word labels often used to describe expressive faces (i.e., Happy, Angry). Words were presented on the screen in boldface, in black ink, font Courier New, font size 24.

Design and Procedure.

Unlike the previous experiments, the maintenance interval was lengthened to 2000 ms. On every trial, during this interval a word appeared in the center of the screen for 100 ms. On half the trials, the word onset 250 ms after the offset of the encoding array (short inter-stimulus-interval, ISI); on remaining trials this ISI was 1000 ms (long ISI). Participants reported the valence of the word (Positive or Negative) as soon as it appeared by pressing one of two keys as quickly and accurately as possible. At the end of the maintenance interval, a single neutral test face appeared in the center of the screen and participants indicated whether that person had been present in the preceding study array or not, as in the previous experiments. Due to the inclusion of the word task, the verbal suppression task was not used.

Each combination of study array face expression (angry, happy), word valence (positive, negative), ISI (short, long), and trial type (match, non-match) was equally probable and presented in a pseudorandom order. Four blocks of 48 trials each were presented in a single session (192 trials in total), so that 24 trials (50% match) contributed to each individual's d' score for each combination of face emotion, word valence, and ISI condition.

Data Analysis

Data from trials in which word task response times (RTs) were less than 200 ms

or longer than the 2000 ms response window (4.51% of trials in total) were eliminated

from the analysis. Percent correct calculations for the word task excluded data from trials

in which the WM response was incorrect; similarly, word task RT analysis excluded data

from trials with errors on either word or WM task or both. Likewise, d', hits, FA, and RT

calculations to index WM performance excluded data from trials with word task errors.

Repeated-measures ANOVAs on word task accuracy, word task RT, and face WM d',

hits, FAs, and RT values were conducted using face emotion (angry, happy), word

valence (negative, positive), and ISI (short, long) as within-subjects factors.

The two different ISI conditions constrained response time on the word valence

task differently. Participants had a maximum of 1750 ms to respond in the short ISI

condition (before the test face appeared) but had only a maximum of 1000 ms in the long

ISI condition. Although the analyses reported here include all trials (regardless of

response window duration), we re-analysed all the results (response time and accuracy on

the word valence task and d' on the faces WM task) using an artificially imposed 1000

ms time window on the short ISI condition. The results of these analyses did not differ

significantly from those reported below.

Results

Word Categorization Task. Accuracy on the word categorization task did not

depend on emotional expression of the WM study array faces, the valence of the word, or

an interaction of these two factors (all p's > .16). Accuracy was, however, 1 percentage

point lower for the short ISI (96%) compared to the long ISI (97%) condition, a

marginally significant effect, $F(1, 21) = 3.46$, $p = .08$, which likely reflects the demands

of consolidating the study faces into WM shortly after they have disappeared from view. ISI did not significantly interact with any other factor (all F's < 1).

Word categorization RT was unaffected by emotional expression of study array faces ($F(1, 21) = 2.09$, $p = .16$). However, word valence ($F(1, 21) = 17.20$, $p < .001$) and ISI ($F(1, 21) = 60.39$, $p < .001$) exerted systematic and significant effects. RTs to categorize word valence were on average 28 ms faster for positive (531.76 ms) than for negative (559.74 ms) words. RTs were also 55 ms faster with the long ISI (518.07 ms) versus short ISI (573.43 ms) condition. However, word valence and ISI did not interact significantly with each other ($F(1, 21) = 2.16$, $p = .16$) or with emotional expression of the study array faces ($F < 1$). Furthermore, the interaction between word valence and face expression was non-significant ($F < 1$), indicating the absence of affective priming by the faces on the word task (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). To confirm the absence of an affective priming effect, we tested a different group of 29 participants using the same stimuli and procedure as for Experiment 3, but asked participants to ignore the WM study array and test faces and to perform the word task only. This mimics the passive viewing of priming stimuli typical in affective priming experiments but includes an atypically long ISI. The RT and accuracy results showed no evidence of affective priming (i.e., non-significant effects of face emotion on word task performance), and fully replicated the word task results reported for Experiment 3, thus confirming that affective priming by faces on words did not occur. This is important because it indicates that emotionally congruent words were not preferentially processed and permits a cleaner interpretation of word effects on WM performance. If affective priming had occurred and congruent words were favoured for processing, this would predict worse performance on

the WM faces task when congruent words were presented. To anticipate, the opposite effect was found.

WM Performance. The main aim of this experiment was to assess whether the irrelevant emotion conveyed by the faces at encoding was maintained during the maintenance interval. If it had been, the valence of the word should have influenced WM for the emotional faces, dependent on the match or mismatch between face and word valence. Indeed, on analysis of d' data this result was obtained (see Figure 4), but only for angry faces, a finding supported by a significant interaction of face emotion and word valence ($F(1, 21) = 7.43$, $p = .01$). Visual WM performance in the angry face condition was significantly better when a negative (2.32 ± 0.18) versus positive (1.93 ± 0.18) word was presented during maintenance ($F(1, 21) = 15.71$, $p = .001$). Visual WM performance on happy face trials was slightly improved when a positive (2.03 ± 0.23) versus negative (1.88 ± 0.26) word was presented, but this effect was small and non-significant ($F < 1$). Note that when positive words were presented, angry face WM fell to a level similar to that found for the happy face WM condition (load 2) of Experiment 1 (1.90 ± 0.27), whereas when negative words were presented WM for angry face identities was comparable to that seen in the angry face (load 2) condition of Experiment 1 (2.50 ± 0.28). Performance in the happy face condition with positive or negative words was comparable to that seen in Experiment 1 (happy, load 2). We also analyzed each word valence condition separately and found that d' was significantly higher for angry versus happy faces when a negative word was presented ($F(1, 21) = 5.67$, $p = .03$) but the corresponding difference for positive words was non-significant ($F < 1$). The main effects of face emotion, word valence, and ISI and all other interactions were non-significant (all

p's > .20). One possible explanation for these findings is that a positive word presented during the maintenance interval reduced general arousal and this had the effect of lowering WM performance for angry faces. However, if this were the case then this reduction in performance should have been seen in the happy face condition also but it was not.

Analysis of Hit rates revealed non-significant main effects of face emotion (F(1, 21) = 2.51, *p* = .13) and ISI ($F < 1$). The main effect of word valence was significant (F(1, 21) = 13.43, *p* = .001), reflecting a greater number of overall hits when the word was negative than positive, a result that could possibly be due to the negative word invoking a higher state of general arousal to the task. However, the same comparison using d' values (and FA rates, reported below) was non-significant, making this finding hard to interpret. There was a marginally significant interaction between face emotion and word valence (F(1, 21) = 3.92, *p* = .06). This interaction reflects significantly greater hits for angry faces when the word was negative (0.84 ± 0.02) versus positive (0.77 ± 0.03) (F(1, 21) = 13.94, *p* = .001), but a non-significant effect of word valence on hits for happy faces ($F <$ 1; negative = 0.79 ± 0.03; positive = 0.78 ± 0.03), mirroring the pattern of results found using d' values. Furthermore, the anger-enhancement effect was observed when the word was negative (F(1, 21) = 7.48, *p* = .01), but the effect of face emotion on hits was non-significant when the word was positive ($F < 1$), as our analysis with d' values also showed. The remaining two- and three-way interactions between face emotion, word valence, and ISI were non-significant (all *p*'s > .45). Hit rates are provided in Table 2.

Analysis of FA rates revealed non-significant main effects of face emotion (F(1, 21) = 1.84, *p* = .19), ISI ($F < 1$), and word valence (F(1, 21) = 2.21, *p* = .15). There was a

significant interaction between face emotion and word valence ($F(1, 21) = 5.90$, $p = .02$) such that emotional congruence produced modestly lower FA rates than emotional incongruence (see Table 2). While this effect of word valence was non-significant in the angry face condition ($F(1, 21) = 1.69$, $p = .21$), FAs in the happy face condition were significantly greater when a negative versus positive word was presented ($F (1, 21) = 5.65$, $p = .03$). The rise in FAs for happy faces when the word was negative may reflect a negativity-induced recognition bias (Johansson et al., 2004; Windmann & Kutas, 2001) that is elicited by the emotional contrast of the intervening negative word.
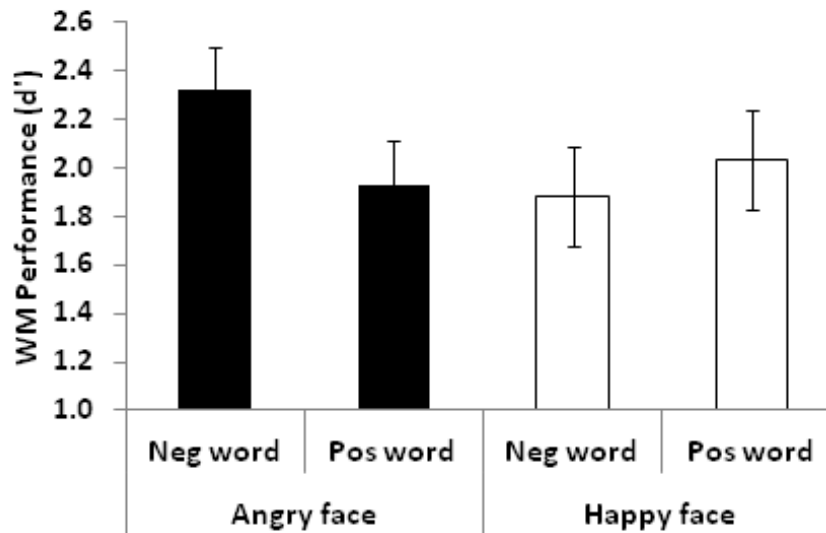

Figure 4. Group mean WM performance (d') scores obtained in Experiment 3 when angry (black bars) or happy (white bars) faces were seen in the study array and positive or negative words were presented during the maintenance interval. The test face was always neutral in expression. Vertical bars indicate ± 1 within-subject standard error.

Table 2. Mean Hits, FA rates, and RTs (ms) for the WM faces task in Experiment 3 as a function of face emotion and word valence. Standard errors are provided in brackets.

	Angry Face		Happy Face	
	Neg Word	Pos Word	Neg Word	Pos Word
Hits	0.84 *(0.02)*	0.77 *(0.03)*	0.79 *(0.03)*	0.78 *(0.03)*
FAs	0.18 *(0.03)*	0.20 *(0.03)*	0.24 *(0.03)*	0.18 *(0.03)*
RTs	1195 *(67)*	1150 *(58)*	1224 *(76)*	1158 *(67)*

A repeated-measures ANOVA on face WM RTs revealed non-significant main effects of face emotion ($F < 1$) and ISI ($F(1, 21) = 1.23$, $p = .28$). Thus, across three experiments RT effects consistently failed to reveal robust effects of face emotion. The main effect of word valence was significant ($F(1, 21) = 8.46$, $p = .008$), reflecting faster RTs to respond to the WM test face overall when the word was positive (1154.01 ms ± 61.04) versus negative (1209.70 ms ± 70.29). This is perhaps a consequence of faster RTs to categorize a positive versus negative word. The interaction between face emotion and word valence was non-significant ($F < 1$) and all remaining two-and three-way interactions were non-significant (all ps > .31). RTs are provided in Table 2.

GENERAL DISCUSSION

In three experiments, we investigated the phenomena of enhanced WM for angry versus happy face identities by manipulating emotional content at three different stages of a simple face memory task. In Experiment 1, we found an anger-specific enhancement of WM when emotional information was manipulated at encoding (study) but remained neutral at retrieval (test). In Experiment 2 we reversed the locus of emotional manipulation, presenting it solely at retrieval, and found no effect of face emotion on WM performance. This pattern of results clearly shows that the presence of a threatening

expression during retrieval is neither necessary nor sufficient to produce anger-specific enhancement of visual WM, and that this effect requires emotional information to be present during encoding.

How might the presence of threatening emotional information at encoding enhance WM for face identity information? The most parsimonious interpretation, based on current related literature, is that angry faces are encoded with greater precision. An emerging theory in the field of visual WM states that variations in WM performance reflect the precision with which stimuli are encoded (Awh, Barton, & Vogel., 2007; Barton, Ester, & Awh, 2009; Bays, Catalao, & Husain, 2009; Bays & Husain, 2008; Fukuda, Awh & Vogel., 2010; Zhang & Luck, 2008). The precision model of WM states that a central pool of resources is flexibly distributed among all items presented in a study array, and items that receive a greater portion of resources are encoded with greater precision (Bays & Husain, 2008; Bays, et al., 2009). For example, decreasing memory accuracy as WM load increases is attributed to the fact that more resources can be allocated per item in small versus large arrays. Sessa and colleagues (2011) proposed that negatively valenced faces are encoded into WM with greater precision than non-threatening faces, referring to evidence that negative information is retained in more detail in long-term memory (e.g., Kensinger, 2007; Kensinger, Garoff-Eaton, & Schacter, 2006, 2007). It is an intriguing and viable possibility that negatively valenced faces receive a greater proportion of resources per item than non-threatening or neutral faces, thus resulting in more detailed WM representations that may in turn enable an appropriate and timely response to threat. For example, if we are confronted by an angry person it would seem beneficial to accurately encode *who* s/he is in order to decide how best to

react – one might respond differently to an angry stranger than an angry friend. On the other hand, an encounter with a happy individual may elicit a similar (benign) response regardless of who is smiling. Supporting the idea of enhanced precision, results from a functional brain imaging study show that enhanced WM for angry versus happy faces is associated with increased neural activity in response to angry versus happy faces in a right hemisphere network of emotion- and face- sensitive regions (Jackson et al., 2008), suggesting that greater resources were allocated to processing angry than happy faces.

Why then might greater resources be allocated to angry than to happy faces? Two possible mechanisms for this emotion-specific boost in processing are selective attention and motivated processing. Selective visual attention, i.e. preferential enhancement of high level visual processing for highly salient or task-relevant visual stimuli, might be a plausible mechanism for anger-specific WM enhancement because a large body of evidence has shown that threat-related stimuli are more effective at capturing attention than neutral or positive stimuli (Eastwood et al.,2003; Feldmann-Wustefeld et al, 2011; Fenske & Eastwood, 2003; Fox & Damjanovic, 2006; Hahn et al., 2006, Horstmann et al., 2006; Huang et al., 2011). This threat bias coupled with evidence that attention enhances perceptual processing (Carassco et al., 2004) suggests that angry faces could have attracted more attention than happy faces during the study periods, leading to greater representational precision during maintenance, and thus to better WM performance. However, there is a substantial problem with this interpretation. Selective attention is primarily needed to prioritize task relevant stimuli over distracting stimuli, so in Experiment 1 task-irrelevant emotional expression information should have been suppressed during encoding in order to boost face identity processing. If, as previous

33

works suggests, angry expressions are more effective at capturing attention than happy

expressions, then the former should have been harder to ignore making identity encoding

even more difficult. This should have led to worse, not better, performance on the angry

face identity WM task.

A more plausible candidate mechanism for the anger-enhancement effect is

'motivated processing'. This refers to the idea that when visual stimuli compete for

access to high level processing (e.g., Desimone & Duncan, 1995), their motivational

value contributes to their competitiveness (Anderson, et al, 2011; Raymond & O'Brien,

2009). Thus, motivational value of a stimulus determines how central resources will be

allocated to it. Stimuli acquire (or may be endowed with) motivational value by being

associated with salient positive or negative outcomes. In this way the brain uses stimuli

and their value associations to predict outcomes, and therefore to plan action and to

modulate the level of processing. Angry expressions signal disapproval and carry an

implicit instruction to alter behavior or "do better" to avoid greater aggression. Happy

expressions, on the other hand, signal approval without implicit instruction for future

actions. Thus, angry faces have greater motivational value and predictive pertinence than

happy faces and therefore attract more central processing resources.

The key question addressed by Experiment 3 was whether face emotion

information is maintained in WM beyond the encoding stage. This experiment replicated

the conditions of Experiment 1 but additionally required a word valence categorization

task during the maintenance interval. If face emotion information is maintained, we

expected to find a significant influence of word valence on WM for the emotional faces

that is dependent on whether faces and words shared the same or different valence.

Whereas analysis of d' values revealed a non-significant effect of word valence on WM for happy faces, word valence significantly modulated performance for angry faces, yielding poorer WM when the intervening word was positive versus negative. Thus, the anger-enhancement effect in WM (as observed in Experiment 1) was only observed when negative words were presented; responding to positive words during the maintenance interval reduced angry face performance to the level found in the happy face conditions and abolished the effect. On more detailed examination of hits and FA rates, we observed that word valence did in fact modulate WM for both angry and happy faces, but in markedly different ways. When study faces were angry, hit rates were significantly greater when a negative versus positive word was presented but FAs were unaffected by word valence. When study faces were happy, hit rates were unaffected by word valence, but FAs were significantly greater when a negative versus positive word was presented.

To account for the isolated influence of word valence on hit rates for angry faces, we propose a 'threat tagging' mechanism in which face identities coupled with an angry expression at encoding are temporarily endowed with a threat association that is sustained during maintenance and galvanized at retrieval when the tagged (but now neutral) identity reappears (i.e., on match trials only). The term threat 'tag' has been used before to describe how long-term memory (LTM) associations between items and their valence can influence attentional biases among clinical populations (Williams, Mathews, and MacLeod, 1996). When items tagged with threat appear in the environment they receive priority processing and are allocated greater resources than items that are not tagged. Richter-Levin and Akirav (2003) proposed that LTM for salient information endowed with an emotional tag is better consolidated, precise, and robust.

While existing theories of emotional tagging relate to LTM, we suggest for the first time here that threat tags, or associations, may modulate processing within the much shorter timescale of working memory, and may be particularly useful during social interaction. Facial expressions of emotion are short-lived and last from only 0.5 to 4 seconds (Ekman, 2003), therefore application of a temporary threat tag to faces in WM would enable an observer to apply and maintain a threat signature that labels a person as angry despite changes in facial expression that might signal otherwise.

Of further note is the question of whether a threat tag is visual or verbal in nature, or takes the form of a more abstract, affective representation (i.e., a feeling). Our finding that face emotion did not significantly modulate accuracy or RTs on the word valence categorization task might suggest that a threat tag is not a verbal label, or at least draws on different resources from those required for the word task (but see the discussion below for an alternative interpretation of these results). However, word valence and facial expression did interact in the context of WM performance, so it is not possible to draw any clear conclusions on this from our current findings. It remains an interesting issue for future research.

Our finding that a positive word categorized during the maintenance of angry faces significantly reduced hit rates (and d' values) relative to a negative word, suggests that the mechanism by which threat becomes durably associated with an encoded face is not infallible to interference. Incongruently valenced intervening information appears to have weakened the threat association and thus reduced its enhancing effects on WM for person identity. The lack of effect of word valence on hit rates for happy faces provides no support for the notion that positive associations are formed or maintained in the same

way as threat associations. Although we did not observe any significant differences in FA rates for happy versus angry faces in Experiment 1 or 2, we did find this in Experiment 3 when an intervening negative word was presented. This effect was driven by a significant increase in FAs for happy faces when the word was negative versus positive, and suggests that some consequence of exposure to the happy face persisted during the maintenance interval. Perhaps the negative word became associated with the originally happy face representation and resulted in a negativity-induced recognition bias (Windmann & Kutas, 2001). Importantly this effect does not appear to be related to WM *per se*.

It is important to note that we do not claim here that angry face representations do not decay at all in WM, but that decay might be more gradual, or have less impact on the precision of representations in WM at the point of retrieval compared to happy face representations. The lack of an anger-enhancement effect when a 9 second maintenance interval was used in previous work (Jackson et al., 2012) indicates clearly that the benefit to WM afforded by anger is lost over longer periods of time. It would be interesting for future study to try to ascertain in more detail the time-course of decay for angry and happy faces and the impact on representational precision.

Finally, with regard to Experiment 3, it is important to consider the perhaps surprising finding that the emotion displayed in the faces at encoding did not impact upon the speed or accuracy with which participants categorized the valence of the intervening word. Does this mean that face valence information was not in fact held in WM during the maintenance / word presentation interval? It has been reported that the contents of WM can guide attention towards shared perceptual features (e.g., colour) present in a

visual search display conducted during the WM maintenance period (Olivers, Meijer, & Theeuwes, 2006; Soto, Heinke, Humphreys, & Blanco, 2005). The notion that information held in WM can enhance attention to other matching or similar information might lead one to predict that when the emotional valence of the faces held in WM matched the valence of the word presented during maintenance, positive/negative categorization of the word should be speeded. However, for this to happen one must assume that the affective contents of WM can influence other valence-related decisions. To the best of our knowledge, there is currently no published research that examines whether the affective, conceptual content of WM (as opposed to featural, perceptual content) can influence concurrent attention to other related stimuli. Furthermore, a significant number of studies did not find an influence of WM on attentional selection for related perceptual information (Downing & Dodds, 2004; Houtkamp & Roelfsema, 2006; Woodman & Luck, 2007).

It is also worth considering whether the lack of influence of face emotion on the word task could be indicative of a threat tag which itself decays within the WM maintenance interval. If a threat tag decayed below a certain threshold by the time a valenced word appeared then this could explain the aforementioned pattern of results. However, this is an unlikely explanation given that the ISI between the offset of the emotional study faces and the onset of the word did not interact with face emotion and/or word valence. In the short ISI condition, the word appeared 250ms after the study faces disappeared and should therefore have proffered the greatest opportunity for face emotion to influence word categorization, compared to the long ISI condition in which the word appeared 1000ms later. A direct assessment of word task performance within just the

short ISI condition confirmed a non-significant interaction between face emotion and word valence for both accuracy and RTs (both F's < 1). Alternatively, since WM accuracy for angry faces was improved when a negative versus positive word was presented, a remaining possibility is that a threat tag does decay rapidly within WM but a congruent negative (not positive) word serves to reactivate the tag during the maintenance period.[1] If this were the case, we may have expected to see a greater boost to angry WM by a negative (but not positive) word at the long versus short ISI, if we assume that a threat tag has suffered greater decay at 1000 ms versus 250 ms after encoding offset. But we do not find support for this. On examination of WM data from only the angry faces condition, the interaction between word valence and ISI was non-significant ($F < 1$ for both d' and Hits), indicating that the time-point at which the word appeared had no measurable influence on WM accuracy as a function of word valence. Nevertheless, this is an interesting concept which requires further investigation to assess properly.

In conclusion, our results reveal some important facets of the interaction between emotional expression and visual WM for faces. Facial expression during a social encounter can change from moment to moment and may not necessarily reflect the valence of a situation at any given point in time. For example, the ability to understand that someone remains angry with you (thus remembering that s/he was angry) when facial expression may indicate otherwise is fundamental to normal human social cognition. Our finding that the influence of an angry expression in visual WM persists despite the disappearance of threat after encoding, and that anger appears to be better sustained throughout the maintenance period than happiness, provides clear evidence that negative

---

[1] We thank an anonymous reviewer for this suggestion.

emotional information confers a potent, lasting impact on face identity memory beyond

initial encounters.

REFERENCES

Anderson, B.A., Laurent, P.A. & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Science,* 108(25), 10367–10371.

Atkinson, R.C. & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. *The Psychology of Learning & Motivation, 2*, 89-195.

Awh, E., Barton, B., & Vogel, E.K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science,* 18 (7) , 622-628.

Bannerman, R.L., Milders, M., & Sahraie, A. (2010). Attentional bias to brief threat-related faces revealed by saccadic eye movements. *Emotion*, 10(5), 733-738.

Barton, B., Ester, E.F. & Awh, E. (2009). Discrete resource allocation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 35(5), 1359-1367.

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

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

Bradley, M.M., & Lang, P.J. (1994). Measuring emotion: The self assessment manikin and the semantic differential. *Journal of Behavioral Therapy and Experimental Psychiatry, 25*(1), 49–59.

Bruce, V. & Young, A. (1986). Understanding face recognition. *British Journal of Psychology,* 77, 305-327.

Buchner, A., Rothermund, K., Wentura, D., & Mehl, B. (2004). Valence of distracter words increases the effects of irrelevant speech on serial recall. *Memory & Cognition, 32*, 722–731.

Carrasco, M., Ling, S. & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, 7(3)*,* 308-313.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioural & Brain Sciences, 24*, 87-185.

Curtis, C.E. & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Science, 7*(9), 415-423.

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

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

Dolcos, F., Iordan, A.D., & Dolcos, S. (2011). Neural correlates of emotion-cognition interactions: a review of evidence from brain imaging investigations. Journal of Cognitive Psychology 23, 669–694.

Dolcos, F. & McCarthy, G. (2006). Brain systems mediating cognitive interference by emotional distraction. Journal of Neuroscience 26, 2072–2079.

Downing, P.E. & Dodds, C.M. (2004). Competition in visual working memory for control of search. *Visual Cognition. 11*, (6) 689-703.

Eastwood, J.D., Smilek, D., & Merikle, P.M. (2003). Negative facial expression captures attention and disrupts performance. *Perception & Psychophysics, 65*(3), 352–358.

Ekman, P. (2003). *Emotions revealed* (2nd ed.). New York: Times Books.

Ekman, P. & Friesen, W.V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Fazio, R.H., Sanbonmatsu, D.M., Powell, M.C., & Kardes, F.R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229–238.

Feldmann-Wüstefeld, T., Schmidt-Daffy, M., & Schubö, A. (2011). Neural evidence for the threat detection advantage: Differential attention allocation to angry and happy faces. *Psychophysiology, 48*(5), 697-707.

Fenske, M.J. & Eastwood, J.D. (2003). Modulation of focused attention by faces expressing emotion: Evidence from flanker tasks. *Emotion, 3,* 327-343.

Fox, E. & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion, 6*(3), 534–539.

Fukuda, K., Awh, E., & Vogel, E.K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology, 20*, 177-182.

Gallegos, D.R. & Tranel, D. (2005). Positive facial affect facilitates the identification of famous faces. *Brain & Language, 93*(3), 338–348.

Galster, M., Kahana, M.J., Wilson, H.R., & Sekuler, R. (2009). Identity modulates short-term memory for facial emotion. *Cognitive and Affective Behavioural Neuroscience, 9*(4), 412-426.

Ganel, T. & Goshen-Gottstein, Y. (2004). Effects of familiarity on the perceptual integrality of the identity and expression of faces: The parallel route hypothesis revisited. *Journal of Experimental Psychology: Human Perception and Performance, 30*(3), 583–597.

Ganel, T., Valyear, K.F., Goshen-Gottstein, Y., & Goodale, M.A. (2005). The involvement of the "fusiform face area" in processing facial expression. *Neuropsychologia, 43*(11), 1645–1654.

Hahn, S., Carlson, C., Singer, S., & Gronlund, S.D. (2006). Ageing and visual search: Automatic and controlled attentional bias to threat faces. *Acta Psychologica, 123*, 312-336.

Henderson, J.M. & Hollingworth, A. (2003). Eye movements and visual memory: detecting changes to saccade targets in scenes. *Perception & Psychophysics*, 65(1), 58-71.

Horstmann, G., Borgstedt, K., & Heumann, M. (2006). Flanker effects with faces may depend on perceptual as well as emotional differences. *Emotion, 6*(1), 28-39.

Houtkamp, R. & Roelfsema, P.R. (2006). The effect of items in working memory on the deployment of attention and the eyes during visual search. *Journal of Experimental Psychology: Human Perception and Performance, 32*(2), 423-442.

Huang, S-L., Chang, Y-C., & Chen, Y-J. (2011). Task-irrelevant angry faces capture attention in visual search while modulated by resources. *Emotion, 11*(3), 544-552.

Jackson, M.C., Linden, D.E.J., & Raymond, J.E. (2012). 'Distracters' do not always distract: Visual working memory for angry faces is enhanced by incidental emotional words. *Frontiers in Psychology, 3*, article 437.

Jackson, M.C. & Raymond, J.E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental psychology: Human Perception & Performance, 34*(3), 556-568.

Jackson, M.C., Wolf, C., Johnston, S.J., Raymond, J.E., & Linden, D.E.J. (2008). Neural correlates of enhanced visual short-term memory for angry faces: An fMRI study. *PLoS ONE, 3*(10), e3536.

Jackson, M.C., Wu, C-Y., Linden, D.E.J, & Raymond, J.E. (2009). Enhanced visual short-term memory for angry faces. *Journal of Experimental Psychology: Human Perception & Performance, 35*(2), 363-374.

Johansson, M., Mecklinger, A., & Treese, A-C. (2004). Recognition memory for emotional and neutral faces: An event-related potential study. *Journal of Cognitive Neuroscience, 16*(10), 1840-1853.

Kastner, S. & Ungerleider, L.G. (2000). Mechanisms of visual attention in the human cortex. *Annual Reviews Neuroscience, 23*, 315-341.

Kaufmann, J.M. & Schweinberger, S.R. (2004). Expression influences the recognition of familiar faces. *Perception, 33*, 399–408.

Kensinger, E.A. (2007). Negative emotion enhances memory accuracy: Behavioral and neuroimaging evidence. *Current Directions in Psychological Science, 16*, 213–218.

Kensinger, E.A., Garoff-Eaton, R.J., & Schacter, D.L. (2006). Memory for specific visual details can be enhanced by negative arousing content. *Journal of Memory and Language, 54*, 99–112.

Kensinger, E.A., Garoff-Eaton, R.J., & Schacter, D.L. (2007). How negative emotion enhances the visual specificity of a memory. *Journal of Cognitive Neuroscience, 19*, 1872–1887.

Klaver, P., Talsma, D., Wijers, A.A., Heinze, H.J., & Mulder, G. (1999). An event-related brain potential correlate of visual short-term memory. *Neuroreport, 10*(10), 2001-2005.

Lee, E., Kang, J.I., Park, I.H., Kim, J-J., & An, S.K. (2008). Is a neutral face really evaluated as being emotionally neutral? *Psychiatry Research, 157*, 77-85.

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

McCollough, A.W., Machizawa, M.G., & Vogel, E.K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex, 43*(1), 77-94.

Mermillod, M., Vermeulen, N., Lundqvist, D., & Niedenthal, P.M. (2009). Neural computation as a tool to differentiate perceptual from emotional processes: The case of anger superiority effect. *Cognition, 110*, 346-357.

Olivers, C.N., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1243–1265.

Raymond, J.E. & O'Brien, J.L. (2009). Selective visual attention and motivation: the consequences of value learning in an attentional blink task. *Psychological Science, 20*(8), 981-988.

Richter-Levin, G. & Akirav, I. (2003). Emotional tagging of memory information – in the search for neural mechanisms. *Brain Research Reviews, 43*, 247-256.

Robitaille, N., Grimault, S., & Jolicoeur, P. (2009). Bilateral parietal and contralateral responses during maintenance of unilaterally encoded objects in visual short-term memory: evidence from magnetoencephalography. *Psychophysiology, 46*(5), 1090-1099.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002) *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

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.

Soto, D., Heinke, D., Humphreys, G.W., & Blanco, M.J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception & Performance, 31*(2), 248-261.

Subramanian, L., Hindle, J.V., Jackson, M.C., & Linden, D.E.J. (2010). Dopamine boosts memory for angry faces in Parkinson's Disease. *Movement Disorders, 25*(16), 2792-2799.

Vogel, E.K. & Machizawa, M.G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*. 428(6984), 748-51.

Vuilleumier, P., Armony, J.L., Driver, J., & Dolan, R.J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron, 30*(3), 829–841.

Vuilleumier, P. & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia, 45*(1), 174–194.

Williams, J.M.G., Mathews, A. & MacLeod, C. (1996). The emotional stroop task and psychopathology. *Psychological Bulletin, 120*(1), 3-24.

Windmann, S. & Kutas, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience, 13*, 577–592.

Woodman, G.F., & Luck, S.J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search? *Journal of Experimental Psychology: Human Perception and Performance, 33*, 363-377.

Yoon, J.H., Curtis, C.E., & D'Esposito, M. (2006). Differential effects of distraction during working memory on delay-period activity in the prefrontal cortex and the visual association cortex. *Neuroimage, 29*(4):1117-26.

Zhang, W. & Luck, S.J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*, 233-235.