Acta Psychologica 219 (2021) 103376

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

# Facing stress: No effect of acute stress at encoding or retrieval on face recognition memory☆

Carey Marr [a,b,*], Conny W.E.M. Quaedflieg [c], Henry Otgaar [a,d], Lorraine Hope [b], Melanie Sauerland [a]

[a] Department of Clinical Psychological Science, Maastricht University, Maastricht, Netherlands
[b] Department of Psychology, University of Portsmouth, Portsmouth, UK
[c] Department of Neuropsychology and Psychopharmacology, Maastricht University, Maastricht, Netherlands
[d] Leuven Institute of Criminology, Catholic University of Leuven, Leuven, Belgium

ARTICLE INFO

*Keywords:*
Stress
Memory
Emotion
Arousal
Eyewitness

ABSTRACT

Eyewitnesses may experience stress during a crime and when attempting to identify the perpetrator subsequently. Laboratory studies can provide insight into how acute stress at encoding and retrieval affects memory performance. However, previous findings exploring this issue have been mixed. Across two preregistered experiments, we examined the effects of stress during encoding and retrieval on face and word recognition performance. We used the Maastricht Acute Stress Test (MAST) to induce stress and verified the success of the stress manipulation with blood pressure measures, salivary cortisol levels, and negative affect scores. To examine differences in stressor timing, participants encoded target faces or words both when confronted with the stressor and during the subsequent cortisol peak and retrieved these stimuli 24 h later. We found neither effects of acute stress on face recognition memory during encoding or retrieval (Experiments 1 and 2), nor effects of encoding stress on word recognition memory (Experiment 2). Bayesian analyses largely provided substantial or strong evidence for the null hypotheses. We emphasize the need for well-powered experiments using contemporary methodology for a more complete understanding of the effect of acute stress on face recognition memory.

## 1. Introduction

The legal system often relies on eyewitnesses to identify perpetrators in the course of criminal investigations. Eyewitnesses may experience stress during a crime and may feel stressed when attempting to identify the perpetrator at the police station (e.g., Bornstein, Hullman, & Miller, 2013; Yuille & Cutshall, 1986). It is known that acute stress at retrieval negatively affects episodic memory performance (Shields, Sazma, McCullough, & Yonelinas, 2017; Wolf, 2017), but these effects are understudied in tasks with more applied eyewitness relevance, such as face recognition. In addition, laboratory experiments examining the effects of acute stress at encoding on episodic memory performance often show mixed findings (e.g., Deffenbacher, Bornstein, Penrod, & McGorty, 2004; Shields et al., 2017; Vogel & Schwabe, 2016). Thus, it remains unclear how acute stress experienced at encoding or retrieval might affect witness face recognition performance.

### 1.1. Effects of acute stress at encoding on memory performance

Whether acute stress during encoding has a negative or positive effect on memory performance seems to depend on the methodology used within a research discipline. In the eyewitness field, most research assessing measures such as recognition accuracy and discriminability reports that acute stress impairs memory performance by (e.g., Davis, Peterson, Wissman, & Slater, 2019; Deffenbacher et al., 2004; Morgan et al., 2004). In contrast, in the fundamental memory field many studies demonstrate that acute stress at encoding can enhance recognition accuracy and discriminability (e.g., Henckens, Hermans, Pu, Joëls, & Fernández, 2009; Vogel & Schwabe, 2016; Zoladz et al., 2011), although meta-analytic results suggest such enhancements occur only under

☆ **Author Note:** This work was supported by a fellowship awarded from the Erasmus Mundus Joint Doctorate Program The House of Legal Psychology (EMJD-LP) with Framework Partnership Agreement (FPA) 2013-0036 and Specific Grant Agreement (SGA) 532473-EM-5-2017-1-NL-ERA MUNDUS-EMJD to Carey Marr, and partly supported with an award from the American Psychology-Law Society Student Grants-in-Aid to Carey Marr.

* Corresponding author at: Universiteitssingel 40, Maastricht 6229 ER, Netherlands.
*E-mail address:* carey.marr@maastrichtuniversity.nl (C. Marr).

https://doi.org/10.1016/j.actpsy.2021.103376
Received 22 February 2021; Received in revised form 6 June 2021; Accepted 14 July 2021
Available online 20 July 2021

certain conditions (e.g., for experiments using stressor-relevant materials and a short delay between stressor and encoding; Shields et al., 2017). These discrepant findings align with expert views on how experienced stress at encoding might affect memory. One recent survey showed that 78% of 36 fundamental memory experts agreed that *Experiencing stress during an event (*i.e.*, at encoding) enhances memory for that event*, whereas only 32% of 37 eyewitness experts agreed with this statement (Marr, Otgaar, Sauerland, Quaedflieg, & Hope, 2020). Methodological differences in the stimuli, induction of stress, stressor verification, stressor timing, and retrieval timing may be the key reasons in explaining these contrasting findings (see Sauerland et al., 2016).

The to-be-remembered stimuli often differ between the fields. First, some evidence suggests that stress effects on memory are stronger for emotional stimuli than neutral stimuli (e.g., Cahill, Gorski, & Le, 2003; Smeets, Otgaar, Candel, & Wolf, 2008), perhaps due to the amygdala's sensitivity to both adrenergic and glucocorticoid stress responses (Joëls, Fernández, & Roozendaal, 2011; McGaugh, 2015). However, a meta-analytic review found no evidence that valence was a moderator of acute stress at encoding on memory performance (Shields et al., 2017). Second, the eyewitness field has more commonly examined encoding stress effects on eyewitness identification or face recognition performance. Such research has often found negative effects of encoding stress on face recognition performance (e.g., Davis et al., 2019; Deffenbacher et al., 2004; Morgan et al., 2004; Pezdek, Abed, & Cormia, 2020). However, in an eyewitness study conducted by Sauerland et al. (2016), participants were stressed using the Maastricht Acute Stress Test (MAST; Smeets et al., 2012) before they witnessed a live theft. Levels of acute stress were confirmed by cortisol measurements. However, this acute stress had no impact on identification performance in a lineup viewed 6–8 days later.

In the fundamental memory field, encoding enhancements have mainly been found with non-facial stimuli such as words, pictures, and slideshows (Domes, Heinrichs, Reichwald, & Hautzinger, 2002; Henckens et al., 2009; Payne et al., 2006). Differences in type of to-be-remembered stimuli may play a role in the contrasting findings, as some research suggests that faces are processed differently than other forms of stimuli (e.g., Diamond & Carey, 1986; Kanwisher & Yovel, 2006; Sato & Yoshikawa, 2013; Woodhead & Baddeley, 1981). Several theories suggest that we may learn and recognize faces differently than we do other types of stimuli. For example, the *face-specificity hypothesis* suggests that faces are distinctively processed in unique areas of the brain (i.e., the fusiform face area) than other stimuli (Kanwisher & Yovel, 2006), and the *expertise hypothesis* proposes that humans are generally experts in face recognition (e.g., Diamond & Carey, 1986). Thus, although there is some evidence to suggest that stress during encoding enhances memory for stimuli like words, static pictures, and slideshows (e.g., Domes et al., 2002; Henckens et al., 2009; Payne et al., 2006; Zoladz et al., 2011; but see Schwabe & Wolf, 2010; Zoladz et al., 2014), it is possible that stress does not affect memory for faces to the same extent. Some experimental results support this idea. For instance, Paul et al. (2016) found that acute stress experienced before a visual discrimination task impaired spatial information but did not affect the discrimination of faces. Wiemers, Sauvage, Schoofs, Hamacher-Dang, and Wolf (2013), using the Trier Social Stress Test (Kirschbaum, Pirke, & Hellhammer, 1993), found that stressed participants outperformed non-stressed participants in face and central object recognition, but not peripheral object recognition. These differences in stimuli may explain at least some of the discrepancies between the eyewitness research and fundamental memory research as fundamental memory experiments most often use words or other non-facial images as stimuli.

Many eyewitness experiments use stressors such as violent videos (Clifford & Hollin, 1981; Cutler, Penrod, & Martens, 1987; Kramer, Buckhout, & Eugenio, 1990), electric shocks (Brigham, Maass, Martinez, & Wittenberger, 1983; Tooley, Brigham, Maas, & Bothwell, 1987), threats of injection (e.g., Maass & Kohnken, 1989; Peters, 1988), and high-fidelity interactive training scenarios (Hope et al., 2016). These stress inductions are typically verified with self-report measures (e.g., Buckhout, Alper, Chern, Silverberg, & Slomovits, 1974; Davis et al., 2019). Other experiments have grouped participants using self-reports about trait stress, state stress, or test anxiety (Bailis & Mueller, 1981; Mueller, Bailis, & Goldstein, 1979; Nowicki, Winograd, & Millard, 1979) and drawn conclusions about stress and memory based on group differences. Such varied methods may elicit different levels of arousal and stress. However, it is difficult to evaluate the efficacy of stress inductions based on self-report data alone, as subjective self-reports of stress do not always correspond to physiological stress responses (Hellhammer & Schubert, 2012). Additionally, methods that increase arousal, thus initiating a noradrenergic response, do not necessarily produce the increases in cortisol that are associated with a physiological stress response (Dickerson & Kemeny, 2004). In contrast, researchers in the fundamental memory field typically use validated laboratory stressors to induce stress (Kirschbaum et al., 1993; Schwabe, Haddad, & Schachinger, 2008; Smeets et al., 2012) with physiological measures serving to verify the stress induction in addition to subjective measures.

Differences in timing of the stress induction and memory encoding may also partly account for the contrasting findings. Stress modulates memory formation and retrieval in a time-dependent manner, closely linked to the temporal action profile of major stress mediators in the brain, in particular noradrenaline and glucocorticoids (Hermans, Henckens, Joëls, & Fernández, 2014; Joëls et al., 2011; Quaedflieg & Schwabe, 2018; Schwabe, Joëls, Roozendaal, Wolf, & Oitzl, 2012). Eyewitness studies mostly examine stress *during* encoding (e.g., most experiments in the Deffenbacher et al. meta-analysis, 2004; Hulse & Memon, 2006; Morgan et al., 2004). Examining this acute stress stage reflects the reality of an eyewitness experience during a crime. In contrast, much of the fundamental memory research conducted on stress during the encoding phase focuses on the delayed stage, that is, when participants encoding information *after* engaging in a stressor (e.g., Wolf, 2012, Exp 2; Zoladz et al., 2011; Quaedflieg, Schwabe, Meyer, & Smeets, 2013), as cortisol peaks about 15–20 min following the stress induction.

The two fields also vary in timing of the encoding and retrieval phases. Many eyewitness experiments examining stress and memory conduct the encoding and retrieval sessions on one day (e.g., Davis et al., 2019; Hulse & Memon, 2006; Valentine & Mesout, 2008). However, a separation between encoding and retrieval is necessary to distinguish stress effects during encoding or consolidation from stress effects during retrieval on memory. If encoding and retrieval take place in one day, then it is impossible to say what memory phase was influenced by the stress. A solution to this issue is to separate encoding and retrieval by at least 24-h, as commonly done in fundamental memory studies (e.g., Shermohammed, Davidow, Somerville, & Murty, 2019; Vogel & Schwabe, 2016; Wolf, 2012, Exp 2; Zoladz et al., 2011, 2014).

To briefly summarize, the fundamental field generally shows that acute encoding stress enhances memory performance (e.g., Henckens et al., 2009; Shields et al., 2017; Vogel & Schwabe, 2016), whereas the eyewitness field suggests that acute stress impairs memory performance (e.g., Davis et al., 2019; Deffenbacher et al., 2004; Morgan et al., 2004). Variations in methodology likely contribute to the discrepant findings regarding the effects of acute stress on memory performance in the eyewitness and fundamental memory fields. As such, when designing studies and interpreting results, it is imperative to consider potential moderators including type of stimuli, type of stress induction and manipulation check, stressor and retrieval timing.

### 1.2. Effects of acute stress at retrieval on memory performance

Stress experienced just prior to retrieval often impairs memory performance (e.g., Quaedflieg & Schwabe, 2018; Shields et al., 2017; Wolf, 2017). This impairment is greatest during the cortisol peak caused by non-genomic actions of glucocorticoids that develop about 15–20 min following the stress induction (de Quervain, Roozendaal, & McGaugh,

1998; Joëls et al., 2011; Joëls & Baram, 2009). Impairments continue as the delayed genomic effects develop, around 60–90 min post-stressor and last for hours (Schwabe & Wolf, 2014; Shields et al., 2017; Wolf, 2017). Accordingly, the specific timing of the stressor is crucial. When the memory test takes place about 20–30 min post-stressor, that is, during the delayed stage (i.e., when stress-induced cortisol increases), results often show memory impairments (Schönfeld, Ackermann, & Schwabe, 2014; Schwabe & Wolf, 2014). When the memory test takes place during the acute stress stage (i.e., just after or during the stress induction and therefore before the stress-induced cortisol peak), such impairments do not regularly occur (e.g., Schwabe & Wolf, 2014). In fact, one study reported that retrieval performance *during* stress was positively associated with the noradrenergic stress response (Schönfeld et al., 2014).

Apart from stressor timing, the type of memory test and valence play a role in the impairing effects of retrieval stress. Stress prior to retrieval impairs both recall (e.g., Kuhlmann, Piel, & Wolf, 2005; Schönfeld et al., 2014; Smeets et al., 2008) and recognition performance (Li, Weerda, Guenzel, Wolf, & Thiel, 2013; Schwabe & Wolf, 2014), with stronger effects for recall than recognition (see Gagnon & Wagner, 2016). Furthermore, retrieval stress tends to elicit larger negative effects for emotional than neutral stimuli (e.g., Kuhlmann et al., 2005; Schönfeld et al., 2014). Thus, the common view is that stress interferes with retrieval performance, though stressor timing, type of memory test, and valence moderate these effects.

Few experiments have investigated the effects of retrieval stress on identification accuracy or face recognition, and findings are mixed. Li et al. (2013) found that stress inducted immediately *prior to* retrieval impaired face recognition sensitivity. However, in a subsequent study the same authors found no statistically significant effects of stress on face recognition accuracy (Li, Weerda, Milde, Wolf, & Thiel, 2014). As far as we are aware, no research has yet examined effects of stress *during* retrieval on face recognition memory. Understanding how experiencing stress during retrieval is likely relevant for applied legal contexts, where witnesses may be stressed when attempting to identify a suspect in a police lineup.

### 1.3. The present experiment

Combining methodology from the eyewitness and fundamental memory fields, the aim of the current two experiments was to examine the effects of acute stress at encoding and retrieval on face and word recognition performance. We induced stress by means of the MAST (Smeets et al., 2012), used physiological measures to verify the stress induction, and inserted a 24–26 h interval between encoding and retrieval. We also tested the effect of stress stage by inserting encoding and retrieval stages both *during* the stress induction (i.e., acute stress stage, when noradrenergic activity ensues) and *after* the stress induction (i.e., delayed stage, when stress-induced cortisol peaks).

In Experiment 1, we examined four between-subjects groups: no stress, stress at encoding, stress at retrieval, and stress at both encoding and retrieval and examined stress stage (acute stress vs. delayed) as a within-subjects measure. Based on models and past findings from the fundamental memory field (e.g., Diamond, Campbell, Park, Halonen, & Zoladz, 2007; Hoscheidt, LaBar, Ryan, Jacobs, & Nadel, 2014; Joëls, Pu, Wiegert, Oitzl, & Krugers, 2006; Vogel & Schwabe, 2016; Wiemers et al., 2013), we predicted that participants who experienced stress during encoding (but not retrieval) would show enhanced face recognition memory compared to the other groups (Hypothesis 1). We also predicted that those who experienced stress during retrieval (but not encoding) would perform the poorest on the face recognition memory task out of the four groups (Hypothesis 2). Furthermore, we expected that the predicted stress effects would be larger during the cortisol peak (i.e., 15 min post-stressor; delayed stage) than before cortisol had peaked (i.e., during stressor; acute stress stage, Hypothesis 3).

Following null results in Experiment 1, Experiment 2 compared participants who experienced encoding stress with those who did not to examine memory performance for face versus word stimuli. We predicted that participants stressed during encoding would show better recognition memory for words than non-stressed participants (Hypothesis 4). Based on our results from Experiment 1 and extant literature concerning the distinctiveness of face processing (e.g., Kanwisher & Yovel, 2006; McGugin, Newton, Gore, & Gauthier, 2014), we expected to obtain similar findings for faces as those in Experiment 1. That is, we hypothesized no statistically significant difference in memory performance for faces between stress conditions (Hypothesis 5). We also examined the role of stress stage (acute stress vs. delayed) though we had no a priori hypotheses for these variables due to the null findings in Experiment 1.

## 2. Method

Both experiments were preregistered on the Open Science Framework (Experiment 1: https://osf.io/k8x5q/?view_only=563a7d16459a47a7aebff54ccd70bf09; Experiment 2: https://osf.io/sqxgb/?view_only=b0564519ca604a17a65b1a34488af187).

### 2.1. Participants

Based on a priori power analyses conducted using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) for an Analysis of Variance (ANOVA) for fixed effects, special, main effects and interactions with 80% power, $\alpha = 0.05$, and a medium-large effect size ($f = 0.26$–$30$[1]; based on past relevant work including Shields et al., 2017), the target sample size for both experiments was $N = 119$.

We recruited participants between the ages of 18 and 35 from the university and local community using posters, handouts/flyers, social media, and lecture visits. Consistent with relevant previous research examining factors that may affect physiological stress reactivity (e.g., Shields, 2020; Strahler, Skoluda, Kappert, & Nater, 2017), we screened for and excluded participants who habitually smoked (>5 cigarettes per day), drank alcohol (>15 drinks per week), or used drugs (more than once per month). Participants were also excluded for a variety of other health reasons (i.e., BMI < 17 and >30, use of medication containing cortisol, recent vaccinations, psychological treatments, cardiovascular problems, or endocrine disorders). Because sex hormones can affect cortisol reactivity (e.g., Kirschbaum, Kudielka, Gaab, Schommer, & Hellhammer, 1999; Kudiekla, Hellhammer, & Wust, 2009; Strahler et al., 2017), females who were not taking a form of hormonal birth control were excluded in order to keep the sample homogeneous. Only white participants were included to avoid influence of the own-race effect on face identification (e.g., Meissner & Brigham, 2001). Additional exclusion criteria for Experiment 2 included participation in Experiment 1 or in another stress-based laboratory study using the MAST within the past month, as past research shows that the MAST can be used repeatedly with no significant signs of habituation or sensitization with intervals of three weeks and one month (Quaedflieg, Meyer, van Ruitenbeek, & Smeets, 2017).

For Experiment 1, 144 participants were tested. Ten participants withdrew during or after Day 1, three experienced a computer program malfunction, and 11 participants indicated having previously seen one or more of the face stimuli and were excluded post-testing. The final sample ($N = 120$) included 41 men and 79 women (age range 18–31 years; $M = 22.04$, $SD = 2.87$). The majority were university students (91.70%) and the rest were not students (8.30%). Most of the students

[1] For Experiment 1, we preregistered that we aimed for a small to medium effect size (between $f = 0.22$ and 0.26). Due to difficulties recruiting eligible participants and a slight change in our analysis plan (as described later), we tested enough participants to provide statistical power for the medium-large effect size of $f = 0.30$, as indicated here.

were completing their Bachelors degrees (75.50%) and the rest were completing a graduate degree (24.50%).

For Experiment 2, 137 participants were tested. Seven participants withdrew on or after Day 1, three experienced a computer program malfunction, and six participants indicated having previously seen one or more of the face stimuli and were excluded post-testing. The final sample ($N = 121$) included 34 men and 87 women (age range 18–34 years; $M = 22.21$, $SD = 2.94$). The majority were university students (94.2%) and the rest were not students (5.80%). Again, most of these students were completing their Bachelors degrees (60.50%), with the rest completing a graduate degree (39.50%).

Participants were asked to follow several rules before Day 1 (Experiments 1 and 2) and Day 2 (Experiment 1). These instructions included not drinking alcohol the night before, getting a full night of sleep, and refraining from eating, drinking anything besides still water, exercising, smoking, or brushing teeth for at least 2 h prior to the session. Participants received either course credit or €20 (Experiment 1) or €15 (Experiment 2) in gift vouchers on completion. This study was approved by the ethical committee of the Faculty of Psychology and Neuroscience at Maastricht University.

### 2.2. Design

Experiment 1 had a 4 (condition: encoding stress vs. retrieval stress vs. stress at encoding and retrieval vs. no stress) x 2 (stress stage: acute stress vs. delayed) mixed design. Stress stage served as a within-subjects factor. Experiment 2 had a 2 (condition: no encoding stress vs. encoding stress) x 2 (stress stage: acute vs. delayed) x 2 (stimulus type: faces vs. words) mixed design. Stress stage and stimulus type were within-subjects factors. For both experiments, participants were semi-randomly assigned to the stress conditions, balancing gender across groups. Our dependent variables included overall accuracy, proportion of hits, proportion of false alarms, sensitivity (*d'*; exploratory for Experiment 1) and response bias indices (*c*; exploratory for Experiments 1 and 2). Sensitivity was calculated by the difference between the z-scores of number of hits and number of false alarms, where the larger *d'*, the better one's performance. Response bias was calculated by the sum of the z-scores of hits and false alarms divided by two, where 0 represents no bias, values less than 0 represent a liberal bias, and values greater than 0 represent a conservative bias.[2] Dependent measures for the manipulation check included negative affect scores, diastolic and systolic blood pressure measurements, and salivary cortisol levels (Experiment 2).

### 2.3. Materials

#### 2.3.1. Stimuli and memory test

*2.3.1.1. Faces (Experiments 1 and 2).* Images were taken from an image database comprised of past and current student and staff volunteers at the Maastricht University. In Experiment 1, we used two color photographs each of 12 males and 12 females as targets (i.e., 48 photographs in total). Targets were all white young adults with a variety of hair colors, hair lengths, hair textures, facial shapes, body types, and eye colors. Photographs of each individual were taken on the same day with no changes in appearance (e.g., haircuts) besides facial expression. Participants viewed two distinct images of each target at encoding, once as a smiling portrait picture and once as a full body picture. In Experiment 2, only the smiling portrait picture was shown (i.e., 24 pictures total). Each image was displayed for 4 s with a 1-s interstimulus interval which allowed for the images to fit within the time-constrained blocks in the (control) MAST procedure. Half of the images (either all males or females) were shown at the acute stress stage and the remaining images were shown at the delayed stage. We counterbalanced gender order and randomized the order of the images in blocks. The recognition test comprised 48 color photographs (24 old, 24 new; 24 male, 24 female) of previously unseen neutral portrait pictures and consisted of a yes/no identification question. We used previously unseen photos at test to ensure that the task measured face recognition as opposed to image recognition (see Burton, 2013). The total number of faces used is comparable to other recent similar work using face recognition tasks (e.g., Davis et al., 2019; Pezdek et al., 2020). Participants provided confidence judgments for all responses on a scale of 0–100% using a sliding bar.[3]

*2.3.1.2. Words (Experiment 2).* Twenty-four negatively valenced words served as targets. Commonly-known negatively valenced nouns were chosen from the Affective Norms for English Words (ANEW; Bradley & Lang, 1999). We used negative words because some research suggests stronger stress effects on memory for emotional stimuli (e.g., Cahill et al., 2003; Joëls et al., 2011; Smeets et al., 2008, but see Shields et al., 2017). Groups of words were balanced for valence, arousal, frequency, and length. The presentation of and retrieval test for words was analogous to the procedure for faces, again with 24 targets and 24 fillers. The order of recognition tests for faces and words were counterbalanced.

#### 2.3.2. Maastricht Acute Stress Test (MAST)

The MAST (Smeets et al., 2012) is a validated laboratory stressor. In this task, participants engage in blocks of hand immersion into ice-cold water (2–4 °C), combined with blocks of socially-evaluated mental arithmetic in front of a critical experimenter who gives negative feedback throughout the task Additionally, participants consent to and are told they are being video recorded for later facial expression analysis. Thus, a second monitor displays a live video of the participant's face to further induce stress, though in reality no recordings are taken. Participants in the no stress conditions were exposed to the control version of this task, which includes hand immersion into room-temperature water (35 °C), basic counting from 1 to 25, and no mention of a video recording. The traditional MAST was slightly varied in these experiments. That is, blocks of stimuli (Experiment 1: faces; Experiment 2: faces and words) were intermixed with the blocks of hand immersion and mental arithmetic. There were four 30 s blocks during the acute stress stage and four 30 s blocks during the delayed stage. This version of the MAST took 16.5 min in total. Fig. 1 depicts a detailed timeline of the procedure.

#### 2.3.3. Affect

The Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) consists of two mood scales containing 10 items that measure positive and negative affect. For each item (e.g., interested, excited, nervous, distressed, etc.), participants indicate to what extent they feel that way at the present moment on a 5-point Likert scale (from *very slightly or not at all* to *extremely*). We measured affect at three timepoints as shown in Fig. 1: during the baseline questionnaires, after the MAST, and at the end of the session after the delayed stage.

#### 2.3.4. Blood pressure

To examine autonomic nervous system activation, we collected systolic and diastolic blood pressure using an Omron Blood Pressure Monitor 705IT (Coleman, Freeman, Steel, & Shennan, 2006). We measured blood pressure eight times throughout the procedure: once during the first saliva sample, once after the MAST anticipation block, twice during the MAST, once just after the MAST, once during the

[2] When proportions of hits or false alarms equaled 0% or 100%, adjusted rates were used for the *d'* and *c* calculations. When 0%, hit or false alarm rates were calculated as $0.5 - n$; When 100%, hit or false alarm rates were calculated as, $(n - 0.5) / n$, where $n$ = number of trials.

[3] Confidence data were not analyzed or reported in this paper.

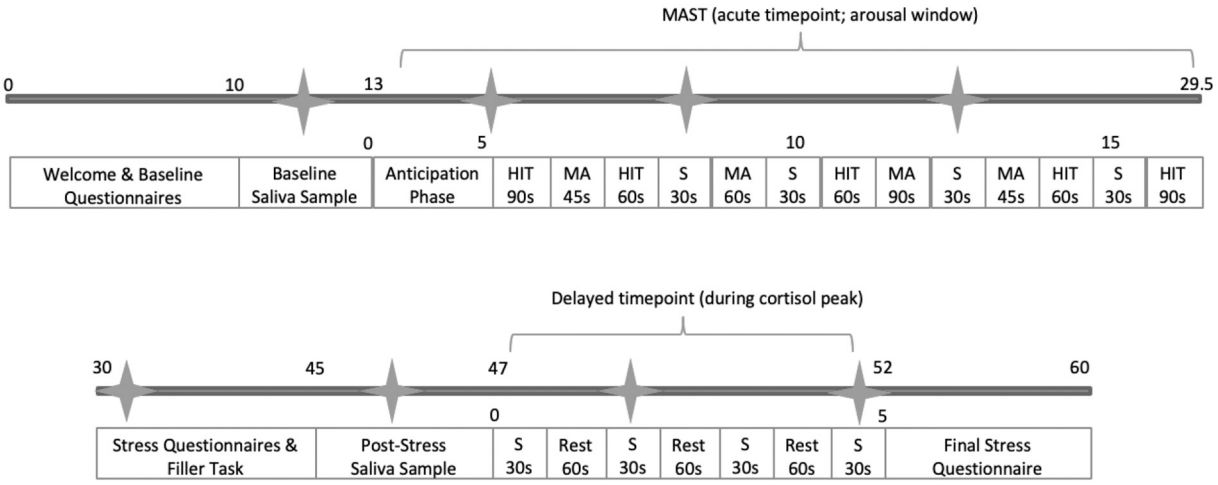


**Fig. 1.** Timeline for Experiment 1 (Day 1 and 2) and Experiment 2 (Day 1)
*Note.* HIT = hand immersion trial. MA = mental arithmetic. S = encoding (Day 1) or recognition (Day 2) of stimuli. Day 2 of Experiment 2 not visually depicted. = blood pressure measurement. MAST = Maastricht Acute Stress Test or control version.

second saliva sample, and twice during the delayed stage (see Fig. 1).

In Experiment 2, the original blood pressure monitor failed, so an updated version (Omron M7 IT HEM-7322 T-E) was used after 11 participants had been tested, although nothing changed procedurally.

#### 2.3.5. Cortisol

We collected saliva samples from participants twice on Day 1 (Experiment 1 and 2) and Day 2 (Experiment 1) with synthetic Salivette (Sarstedt®, Etten-Leur, the Netherlands) devices. In the absence of funding for sample processing, we have not analyzed the samples from Experiment 1. The saliva samples from Experiment 2 were frozen and stored at −20 °C until analysis. After thawing, Salivettes were centrifuged at 3000 rpm for 5 min, which resulted in a clear supernatant of low viscosity. Salivary concentrations were measured using commercially available chemiluminescence immunoassay with high sensitivity (IBL International, Hamburg, Germany). The intra- and inter-assay coefficients for cortisol were both below 9%.

### 2.4. Procedure

We tested participants in both experiments between 12:00 pm and 6:00 pm to account for diurnal cortisol levels (Shields, 2020). After giving consent, participants completed the PANAS and demographic questions as well as other filler questionnaires unrelated to this study. Participants were aware that they would be asked to recognize the stimuli during the follow-up session. After 10 min, participants provided a saliva sample and engaged in the (control) MAST, while viewing blocks of faces (Experiment 1) or blocks of faces or words (Experiment 2). This constituted the acute stress stage of face encoding. Next, participants again completed the PANAS before participating in a filler task (Tetris). Once 15 min had passed since the end of the (control) MAST (to await the expected cortisol response; Smeets et al., 2012), participants provided another saliva sample and completed the delayed stage of face encoding, consisting of blocks of faces (Experiment 1) or faces and words (Experiment 2) in 30 s chunks which mimicked encoding at the acute stress stage. Following this, participants once again completed the PANAS. The Day 1 session took 1 h to complete.

Participants returned to the lab 24–26 h later. For Experiment 1, the procedure on Day 2 was identical to Day 1 with the exception that instead of encoding faces during and after the (control) MAST participants took part in a face recognition test at these times. For Experiment 2, the Day 2 session did not involve engagement with the (control) MAST but instead consisted only of a recognition test which took 15 min to complete. Following the recognition test, participants received a debriefing and compensation. Fig. 1 shows an overview of the procedure timeline.

### 2.5. Data analysis

To verify the stress manipulations, we conducted mixed ANOVAs on the effect of stress condition (i.e., stress vs. no stress) on self-reported negative affect scores and systolic and diastolic blood pressure across each testing day (i.e., timing: three levels for negative affect; eight levels for blood pressure). For Experiment 2, we also conducted a mixed ANOVA on the effect of stress condition on salivary cortisol pre- and post-stressor. Our main analyses for Experiment 1 differed to those that were pre-registered. These deviations were decided upon prior to analyzing the data. Specifically, rather than conducting two separate 2 (stress at encoding: yes vs. no) x 2 (stress at retrieval: yes vs. no) for each testing stage, we conducted a mixed ANOVA with condition (no stress vs. encoding stress vs. retrieval stress vs. both) as between subjects and stress stage (acute stress vs. delayed) as within subjects factor. We made this decision in order to compare the stages as past research has emphasized that stress effects on memory are time sensitive (e.g., Quaedflieg & Schwabe, 2018; Schwabe et al., 2012). Distinctly examining the four stress conditions in one analysis also allowed for a clearer test of our hypotheses, which placed emphasis on the encoding only and the retrieval only groups. Thus, for both experiments, we conducted one-way mixed ANOVAs to analyze the effects of condition on overall accuracy, proportion of hits, proportion of false alarms, sensitivity, and response bias. For all tests, when Mauchly's test indicated that the assumption of sphericity was violated, we corrected the degrees of freedom using Greenhouse-Geisser estimates of sphericity.

## 3. Results

### 3.1. Manipulation checks

Table 1 provides an overview of the main inferential statistics for the manipulation checks across both experiments.

#### 3.1.1. Negative affect

Higher scores on the negative affect portion of the PANAS reflect higher self-reported negative affect. In Experiment 1 on Day 2, negative affect scores were missing for one participant ($N = 119$). Across both experiments, stress differentially affected negative affect scores

**Table 1**
Inferential Statistics for Stress x Timing Interactions and Simple Main Effects at Baseline and Post-Stressor across Experiments.

| | | Experiment 1 Day 1 | | | Experiment 1 Day 2 | | | Experiment 2 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Stress x Timing Interaction | Simple Effects of Stress | | Stress x Timing Interaction | Simple Effects of Stress | | Stress x Timing Interaction | Simple Effects of Stress | |
| | | | Baseline | Post-Stressor | | Baseline | Post-Stressor | | Baseline | Post-Stressor |
| Negative affect | $F$ | **34.703** | 0.004 | **39.909** | **21.130** | 3.932 | **31.712** | **29.160** | 0.740 | **41.454** |
| | $p$ | **<0.001** | 0.948 | **<0.001** | **<0.001** | 0.050 | **<0.001** | **<0.001** | 0.391 | **<0.001** |
| | $\eta_p^2$ | **0.227** | <0.001 | **0.213** | **0.153** | 0.033 | **0.213** | **0.197** | 0.006 | **0.258** |
| Systolic blood pressure | $F$ | **12.787** | 0.162 | **19.102** | **12.393** | 0.353 | **14.426** | **10.939** | 0.578 | **27.360** |
| | $p$ | **<0.001** | 0.688 | **<0.001** | **<0.001** | 0.554 | **<0.001** | **<0.001** | 0.449 | **<0.001** |
| | $\eta_p^2$ | **0.098** | 0.001 | **0.139** | **0.095** | 0.003 | **0.109** | **0.085** | 0.005 | **0.188** |
| Diastolic blood pressure | $F$ | **10.776** | 0.010 | **20.838** | **6.972** | 0.658 | **9.479** | **14.858** | 0.324 | **29.262** |
| | $p$ | **<0.001** | 0.921 | **<0.001** | **<0.001** | 0.419 | **0.003** | **<0.001** | 0.570 | **<0.001** |
| | $\eta_p^2$ | **0.084** | <0.001 | **0.150** | **0.056** | 0.006 | **0.074** | **0.110** | 0.003 | **0.199** |
| Salivary cortisol | $F$ | – | – | – | – | – | **–** | **66.535** | 0.714 | **31.229** |
| | $p$ | – | – | – | – | – | – | **<0.001** | 0.400 | **<0.001** |
| | $\eta_p^2$ | – | – | – | – | – | – | **0.359** | 0.006 | **0.208** |

*Note*. Experiment 1, Day 1: $N = 120$. Experiment 1, Day 2: $N = 119/120$. Experiment 2: $N = 120/121$. Findings in bold are statistically significant at the $p < .05$ level.

depending on the timing (see Table 1). The pattern of results was similar in each analysis, with follow-up tests revealing no statistically significant differences between groups at baseline (all $ps \geq 0.050$, see Table 1). However, right after the stressor (i.e., MAST), stressed participants reported statistically significantly higher levels of negative affect than non-stressed participants (all $ps \leq 0.001$, see Table 1). Thus, negative affect scores in both experiments subjectively confirm the stress induction. Fig. 2 displays changes in negative affect scores across time in both experiments.

#### 3.1.2. Blood pressure

In Experiment 2, blood pressure measurements were missing for one participant ($N = 120$). In both experiments, stress differentially affected both systolic and diastolic blood pressure depending on the timing (see Table 1). Follow-up tests revealed no statistically significant differences between stress conditions at baseline (all $ps \geq 0.419$, see Table 1). However, as expected, during the stressor (i.e., MAST 1), the stress groups showed statistically significantly higher blood pressure than the non-stressed group (all $ps \leq 0.003$, see Table 1). Thus, blood pressure measures in both experiments confirm physiological arousal. Fig. 3 shows changes in blood pressure measurements across time on each testing day.

#### 3.1.3. Salivary cortisol (Experiment 2)

Stress differentially affected salivary cortisol levels as a function of timing (see Table 1). Follow-up tests revealed no statistically significant differences in salivary cortisol between stress conditions at baseline ($p = .400$, see Table 1). As expected, after the stressor, the stress group showed statistically significantly higher salivary cortisol levels than the non-stressed group, ($p < .001$, see Table 1). Thus, the salivary cortisol levels confirm the acute stress induction. Fig. 4 shows changes in salivary cortisol levels before and after the stressor. We also exploratorily grouped cortisol responders vs. non-responders, with 62.30% of the stress group showing a high cortisol response and the other 37.70% showing a low cortisol response (1.5 nmol/l increase, Miller, Plessow, Kirschbaum, & Stalder, 2013). However, this exploratory analysis examining high responders, low responders, and participants who were not stressed revealed no statistically significant differences of acute stress on any memory measure (all $ps > 0.050$; see Table A in supplementary materials), and thus, no further analyses on these subgroups were performed or reported.

### 3.2. Effects of acute stress on memory performance

Tables 2-4 provide a summary of the inferential and descriptive statistics for all dependent variables regarding the main analyses from each experiment.

#### 3.2.1. Experiment 1

For face recognition, we found no statistically significant effects of condition, stress stage, or interactions between these factors on overall accuracy, proportion of hits, or proportion of false alarms[4,5] Thus, Hypotheses 1, 2, and 3 were not supported. We also examined sensitivity ($d'$) and response bias ($c$) as exploratory (i.e., non-preregistered) outcome variables. There were no statistically significant effects for sensitivity, $ps \geq 0.060$. For response bias, the main effect of stress stage was statistically significant. Fig. 5 illustrates that responding was more liberal at the acute stress stage than at the delayed stage although participants in general were liberal rather than conservative in their responding.

We also examined the data using Bayesian ANOVAs with JASP version 0.13.1 (JASP Team, 2020). We adopted a weakly informative prior by setting the $r$ scale fixed effect at a default value of 0.5. Results were similar to those reported above, with only the evidence regarding response bias ($c$) pointing away from the null hypothesis. Specifically, for the main effect of stress stage on response bias, there was very strong evidence for the alternative hypothesis ($BF_{10} = 37.037$; Jarosz & Wiley, 2014; Jeffreys, 1961; Raftery, 1995). All other estimated BFs examining main and interaction effects of the dependent variables suggested the data were in favor of the null hypothesis, with evidence ranging from anecdotal to strong (i.e., from $BF_{01} = 1.340$ to $BF_{01} = 13.356$, see Table D in supplementary materials).

#### 3.2.2. Experiment 2

Memory measure data were missing for one participant, and thus they were excluded from the following analyses ($N = 120$). Failing to support Hypothesis 4, we found no evidence for effects of acute stress on word recognition performance. However, supporting Hypothesis 5, we again found no statistically significant evidence that acute stress affected face recognition memory. The main effect of condition was non-significant for all dependent measures. A main effect of stimulus type indicated a higher false alarm rate for words than faces. Statistically significant interactions between stimulus type and stress stage on hit

[4] As preregistered, we also conducted the analyses after removing outliers (scores ≥2.5 *SD*s from the mean; see Table B in supplementary materials). These analyses (*N*s = 112/118) returned analogous patterns of results.

[5] Additionally, we analyzed data with guesses (confidence at 50% or below) removed (see Table C in supplementary materials). No statistically significant interaction or main effects emerged for any of the outcome variables. Thus, the main effect of stress stage on response bias was no longer statistically significant although responding was still in general liberal rather than conservative.

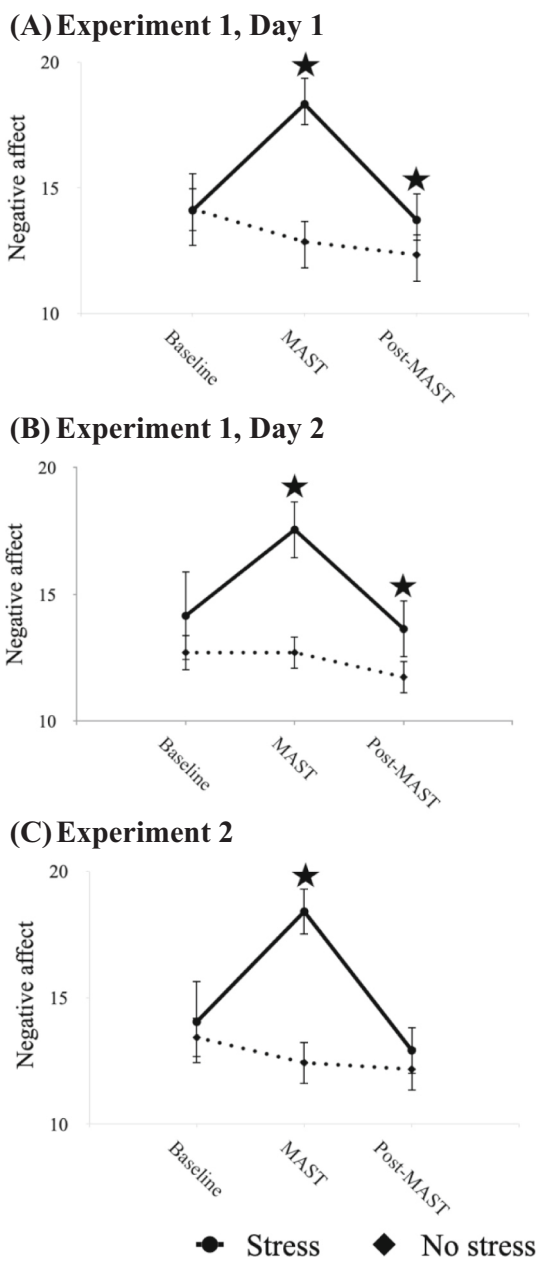


**Fig. 2.** Negative affect scores over time in Experiments 1 and 2 across stress conditions.

*Note.* Possible negative affect scores range between 10 and 50. Negative affect scores are from the Positive and Negative Affect Schedule. MAST = Maastricht Acute Stress Test. Error bars = 95% confidence intervals. = $p < .05$.

rate and response bias modified main effects of stimulus type. Specifically, words elicited a higher hit rate and a more conservative response bias than faces at both the acute stress stage, hits: $F = 170.936$, $p < .001$, $\eta_p^2 = 0.592$; *c:* $F = 246.715$, $p < .001$, $\eta_p^2 = 0.676$, and the delayed stage, hits: $F = 56.189$, $p < .001$, $\eta_p^2 = 0.323$; *c*: $F = 129.236$, $p < .001$, $\eta_p^2 = 0.523$. Additionally, the stimulus type by stress stage interaction was statistically significant for overall accuracy and sensitivity. Follow-up tests indicated higher overall accuracy for words as opposed to faces during the acute stress stage, $F = 10.350$, $p = .002$, $\eta_p^2 = 0.081$, but not the delayed stage, $F = 0.572$, $p = .451$, $\eta_p^2 = 0.005$. Similarly, *d'* scores for words as opposed to faces were higher during the acute stress stage, $F = 9.154$, $p = .003$, $\eta_p^2 = 0.072$, but not the delayed stage, $F = 0.994$, $p = .321$, $\eta_p^2 = 0.008$.[6,7]

Bayesian ANOVAs revealed decisive evidence for the alternative hypothesis for the main effect of stimulus on proportion of hits, proportion of false alarms, and response bias ($BF_{10} = 8.953e36$, $BF_{10} = 4.929e35$, $BF_{10} = 3.789e57$, respectively). For the interaction between stimulus type and stress stage, there was substantial evidence for the alternative hypothesis for overall accuracy ($BF_{10} = 8.065$) and sensitivity ($BF_{10} = 7.519$), strong evidence for response bias ($BF_{10} = 24.390$), and decisive evidence for proportion of hits ($BF_{10} = 1037.883$). All other results showed more evidence towards the null hypothesis, ranging from anecdotal to substantial evidence (i.e., from $BF_{01} = 1.408$ to $BF_{01} = 9.524$). We report all BFs and related information in Table G in supplementary materials.

## 4. General discussion

Across two preregistered experiments, we applied contemporary methodology from the eyewitness and fundamental memory fields to study the effects of acute stress at encoding and retrieval on face recognition performance. We induced stress with a validated laboratory stressor and verified the success of the stress induction physiologically in addition to standard subjective verification. To allow for a separation of stress effects at each memory phase, we inserted a 24 to 26-h interval between encoding and retrieval. Contrary to our hypotheses and previous work both in the eyewitness and the fundamental memory fields, we found no effect of acute stress on face recognition memory during encoding or retrieval and no effects of encoding stress on word recognition memory. The interpretation of the findings as support of the null hypothesis were largely supported by substantial or strong evidence relying on Bayesian analyses (Jarosz & Wiley, 2014; Jeffreys, 1961; Raftery, 1995).

Our findings contrast with previous fundamental work on stress and memory showing memory enhancement for faces encoded during a stressor (Wiemers et al., 2013) and memory impairment for faces retrieved after exposure to a stressor (Li et al., 2013). At the same time, our findings also contrast with previous eyewitness research suggesting negative effects of acute encoding stress on face recognition performance (Davis et al., 2019; Deffenbacher et al., 2004; Pezdek et al., 2020) although differences in methodology may explain these conflicting results. Indeed, the most relevant eyewitness experiment, which examined acute stress effects on identification performance using similarly robust methodology and a lineup identification task, also found no effects of acute encoding stress on memory performance (Sauerland et al., 2016). These results coupled with our current null findings call into question the robustness and generalizability of past findings regarding the effects of acute stress on face recognition memory.

The absence of effects of acute stress on word recognition is even more surprising, given the comprehensive literature supporting this idea (e.g., Abercrombie, Kalin, Thurow, Rosenkranz, & Davidson, 2003; Domes et al., 2002; Schwabe, Bohringer, Chatterjee, & Schachinger, 2008; Zoladz et al., 2011). However, many such experiments showed

[6] Analyses after removing outliers (scores ≥ 2.5 *SDs* from the mean; see Table E in supplementary materials; *Ns* = 113/117) returned analogous patterns of results.

[7] Analyses after removing guesses (responses with ratings of less than 51% confidence; see Table F in supplementary materials; $N = 120$) showed the same pattern of results. In addition, main effects of stress stage emerged for two of the outcome variables, false alarms and *d'*. There was a greater proportion of false alarms ($M = 0.337$, $SE = 0.016$) and during the acute stress stage than during the delayed stage ($M = 0.302$, $SE = 0.016$). The main effect of stress stage for *d'* was modified by the interaction between stimulus type and stress stage.

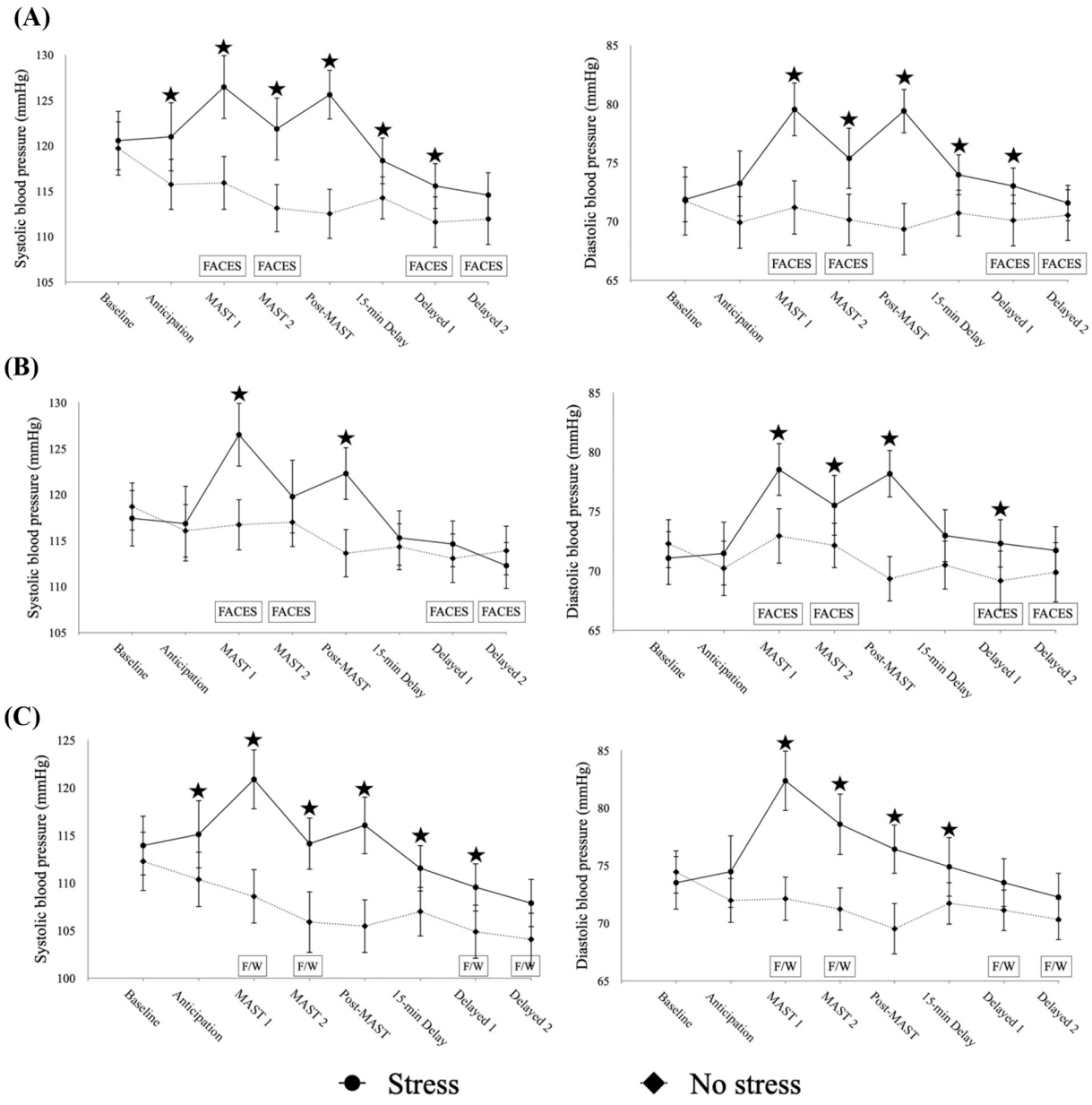


**Fig. 3.** Systolic and diastolic blood pressure over time in Experiments 1 and 2 across stress conditions.
*Note.* (A) = Experiment 1, Day 1. (B) = Experiment 1, Day 2. (C) = Experiment 2, Day 1. mmHg = millimeters of mercury. MAST = Maastricht Acute Stress Test (acute stress stage). Delayed = delayed stage. Faces and F/W = times when faces (and words) were encoded or recognized. Error bars = 95% confidence intervals. = $p < .05$.

enhancement effects only under certain conditions. For example, some experiments showed enhancements only for positively valenced (but not neutral or negative words; Zoladz et al., 2011) or neutral words (but not positive or negative words; Schwabe, Bohringer, et al., 2008). Other experiments showed effects for participants classified as high cortisol responders (e.g., Domes et al., 2002), but not for the stress group as a whole. However, notably, results from Experiment 2 do mirror some other fundamental work demonstrating a lack of effect of cortisone administration (de Quervain, Roozendaal, Nitsch, McGaugh, & Hock, 2000) or acute encoding stress (Domes, Heinrichs, Rimmele, Reichwald, & Hautzinger, 2004) on word recognition performance, suggesting that such findings are not entirely atypical.

We also tested the effect of encoding stress both during and after the stress induction to examine potentially different effects of acute stress on stress stage (e.g., Quaedflieg & Schwabe, 2018; Shields et al., 2017). Effects during the stress induction reflect the reality of an eyewitness experience during a crime, whereas effects after the stress induction reflect the time window when cortisol exerts the strongest effect on memory. One previous study found recognition enhancements for items experienced during a stressor and 41–65 min post-stressor (Vogel & Schwabe, 2016). Although our stress induction was successful, we did not see any memory benefit for stressed participants in either stress stage. Timing differences might partially explain this discrepancy: the delayed stage in our study occurred around 34–39 min post-stressor onset, slightly earlier than the 41–65 min period specified in this earlier experiment. We designed our study to align with the anticipated cortisol peak (Joëls et al., 2011; Joëls & Baram, 2009; Quaedflieg & Schwabe, 2018). Still, other research also illustrates conflicting results at various stress stages, such as showing that encoding during a stressor impaired recognition memory (Schwabe & Wolf, 2014) or demonstrating that encoding immediately post-stressor, but not 30-minute post-stressor, enhanced recognition of positive words (Zoladz et al.,

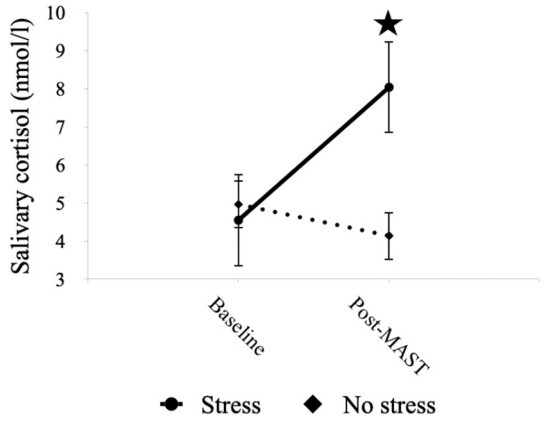


**Fig. 4.** Salivary Cortisol Level Pre- and Post- MAST in Experiment 2 Across Stress Conditions
*Note.* MAST = Maastricht Acute Stress Test. Error bars = 95% confidence intervals. ⋆ = $p < .05$.

2011). Placing our null results alongside these mixed findings highlights the need for further investigations of the specific timeline regarding effects of acute encoding stress on recognition memory.

Perhaps the most promising explanation for the lack of acute stress effects at both encoding and retrieval on recognition performance is to consider the type of memory test used. A meta-analysis examining cortisol administration effects on encoding and retrieval showed that effect sizes for recognition memory were smaller than effect sizes for free or cued recall memory (Het, Ramlow, & Wolf, 2005). Likewise, for retrieval stress, the effect of stress or cortisol seems to differ as a function of memory test type, with stronger negative effects for recall than recognition performance (de Quervain et al., 2000, 2003; Gagnon & Wagner, 2016; Wolf, 2017). The underlying mechanism could be that recognition memory relies on familiarity, whereas free recall requires recollection. Recall is associated with greater hippocampal dependency than recognition (Gagnon & Wagner, 2016). As a result, acute stress, which directly affects the hippocampus, may more strongly impair free recall, while sparing familiarity-based judgments. Thus, the type of memory test may influence encoding and retrieval stress effects on memory performance, perhaps explaining the null results in the current experiments. Directly comparing recognition and recall in future experiments could reveal valuable insights about the extent of effects of acute stress on different kinds of memory performance.

The current experiments also have some limitations. In Experiment 2, we found statistically significant cortisol increases in stressed participants. Examining salivary cortisol effects in Experiment 1 may also have provided a fuller picture of cortisol-related stress effects. For example, some research suggests negative effects of acute retrieval stress on memory emerge only in high cortisol responders (e.g., Buchanan & Tranel, 2008; Buchanan, Tranel, & Adolphs, 2006; Schönfeld et al., 2014). Therefore, examining cortisol responder groups may have helped clarify the null effects of retrieval stress on recognition in Experiment 1. Further, in laboratory experiments—where stress inductions are constrained by ethical considerations—results reflect mild to medium levels of acute stress, rather than severe levels of stress that may be present in certain extreme criminal contexts. The face recognition tasks in Experiments 1 and 2 also differed from one another in task difficulty. Unfamiliar face recognition is already a difficult task (e.g., Hancock, Bruce, & Burton, 2000), and contrasts with other types of recognition tasks, in

**Table 2**
Memory Performance Measures (Means and SDs) as a Function of Condition and Stress Stage and Inferential Statistics for Experiment 1.

| | No stress | | Encoding stress | | Retrieval stress | | Encoding and retrieval stress | | | Main effect stress stage | Main effect condition | Interaction stress stage x condition |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Stress stage | Acute | Delayed | Acute | Delayed | Acute | Delayed | Acute | Delayed | | | | |
| Overall accuracy | 0.858 (0.109) | 0.844 (0.104) | 0.794 (0.109) | 0.796 (0.104) | 0.840 (0.109) | 0.833 (0.104) | 0.799 (0.109) | 0.811 (0.104) | $F$ | 0.0370 | 2.2960 | 0.3970 |
| | | | | | | | | | $p$ | .848 | .0810 | .7560 |
| | | | | | | | | | $\eta_p^2$ | < 0.001 | .056 | .010 |
| Hits (%) | 0.800 (0.164) | 0.789 (0.175) | 0.733 (0.164) | 0.736 (0.175) | 0.772 (0.164) | 0.775 (0.175) | 0.697 (0.164) | 0.756 (0.175) | $F$ | 0.7130 | 1.4280 | 0.9710 |
| | | | | | | | | | $p$ | .4000 | .2380 | .4090 |
| | | | | | | | | | $\eta_p^2$ | .006 | .036 | .024 |
| False alarms (%) | 0.083 (0.126) | 0.100 (0.137) | 0.144 (0.126) | 0.144 (0.137) | 0.092 (0.126) | 0.108 (0.137) | 0.100 (0.126) | 0.133 (0.137) | $F$ | 2.3320 | 1.1760 | 0.3890 |
| | | | | | | | | | $p$ | .1290 | .3220 | .7610 |
| | | | | | | | | | $\eta_p^2$ | .020 | .030 | .010 |
| *d'* | 1.842 (0.887) | 2.228 (0.676) | 1.844 (0.887) | 1.882 (0.676) | 2.076 (0.887) | 2.141 (0.767) | 1.783 (0.887) | 1.955 (0.767) | $F$ | 3.5980 | 0.9830 | 0.8250 |
| | | | | | | | | | $p$ | .0600 | .4030 | .4830 |
| | | | | | | | | | $\eta_p^2$ | .030 | .025 | .021 |
| *c* | −0.468 (0.394) | −0.203 (0.400) | −0.228 (0.394) | −0.224 (0.400) | −0.313 (0.394) | −0.215 (0.400) | −0.433 (0.394) | −0.203 (0.400) | $F$ | **12.3080** | 0.7270 | 2.0010 |
| | | | | | | | | | $p$ | **.0010** | .5380 | .1180 |
| | | | | | | | | | $\eta_p^2$ | **.096** | .018 | .049 |

*Note.* $N = 120$. *d'* = sensitivity. *c* = response bias. Acute = acute stress stage. Delayed = delayed stage. Numbers in parentheses represent standard deviation. Findings that are statistically significant at the $p < .05$ level are in bold. A post-hoc Bonferroni test indicated more liberal responding at the acute stress stage ($M = -0.360$, $SE = 0.036$) than at the delayed stage ($M = -0.211$, $SE = 0.037$).

**Table 3**
Memory performance measures for words and faces as a function of stress condition and stress stage in Experiment 2.

| | Words | | | | Faces | | | |
|---|---|---|---|---|---|---|---|---|
| | Acute | | Delayed | | Acute | | Delayed | |
| | Stress | No stress | Stress | No stress | Stress | No stress | Stress | No stress |
| Overall accuracy | 0.694 (0.102) | 0.682 (0.100) | 0.676 (0.109) | 0.672 (0.115) | 0.639 (0.102) | 0.650 (0.100) | 0.703 (0.109) | 0.664 (0.115) |
| Hits (%) | 0.777 (0.156) | 0.792 (0.154) | 0.732 (0.172) | 0.715 (0.177) | 0.496 (0.219) | 0.487 (0.215) | 0.586 (0.195) | 0.534 (0.200) |
| False alarms (%) | 0.396 (0.180) | 0.398 (0.177) | 0.372 (0.164) | 0.370 (0.161) | 0.212 (0.156) | 0.179 (0.154) | 0.180 (0.164) | 0.206 (0.169) |
| *d'* | 1.151 (0.656) | 1.193 (0.653) | 1.077 (0.656) | 1.023 (0.653) | 0.878 (0.719) | 0.970 (0.714) | 1.253 (0.719) | 1.012 (0.714) |
| *c* | 0.276 (0.422) | 0.308 (0.422) | 0.175 (0.422) | 0.138 (0.422) | −0.457 (0.476) | −0.544 (0.476) | −0.390 (0.469) | −0.434 (0.469) |

*Note.* $N = 120$. *d'* = sensitivity. *c* = response bias. Acute = acute stress stage. Delayed = delayed stage. Numbers in parentheses represent standard deviation.

**Table 4**
Inferential statistics for effects of condition, stimulus type, and stress stage on memory performance measures in Experiment 2.

| | | Main effect condition | Main effect stimulus | Main effect stress stage | Interaction condition x stimulus | Interaction condition x stress stage | Interaction stimulus x stress stage | Interaction condition x stimulus x stress stage |
|---|---|---|---|---|---|---|---|---|
| Overall accuracy | *F* | 0.605 | 3.493 | 1.410 | 0.264 | 1.812 | **12.939** | 2.623 |
| | *p* | 0.438 | 0.064 | 0.238 | 0.609 | 0.181 | **<0.001** | 0.108 |
| | $\eta_p^2$ | 0.005 | 0.029 | 0.012 | 0.002 | 0.015 | **0.099** | 0.022 |
| Hits (%) | *F* | 0.502 | **171.243** | 0.068 | 0.679 | 2.012 | **22.778** | 0.041 |
| | *p* | 0.480 | **<0.001** | 0.795 | 0.405 | 0.159 | **<0.001** | 0.841 |
| | $\eta_p^2$ | 0.004 | **0.592** | 0.001 | 0.006 | 0.017 | **0.162** | <0.001 |
| False alarms (%) | *F* | 0.005 | **164.373** | 1.337 | 0.014 | 1.209 | 1.157 | 1.904 |
| | *p* | 0.943 | **<0.001** | 0.274 | 0.905 | 0.274 | 0.284 | 0.170 |
| | $\eta_p^2$ | <0.001 | **0.582** | 0.010 | <0.001 | 0.010 | 0.010 | 0.016 |
| *d'* | *F* | 0.289 | 1.622 | 0.584 | 0.285 | 3.595 | **10.459** | 1.348 |
| | *p* | 0.592 | 0.205 | 0.446 | 0.594 | 0.060 | **0.002** | 0.248 |
| | $\eta_p^2$ | 0.002 | 0.014 | 0.005 | 0.002 | 0.030 | **0.081** | 0.011 |
| *c* | *F* | 0.356 | **291.182** | 0.637 | 0.634 | 0.048 | **13.465** | 0.822 |
| | *p* | 0.552 | **<0.001** | 0.426 | 0.427 | 0.826 | **<0.001** | 0.366 |
| | $\eta_p^2$ | 0.003 | **0.712** | 0.005 | 0.005 | <0.001 | **0.102** | 0.007 |

*Note.* $N = 120$. *d'* = sensitivity. *c* = response bias. Acute = acute stress stage. Delayed = delayed stage. Findings that are statistically significant at the $p < .05$ level are in bold.

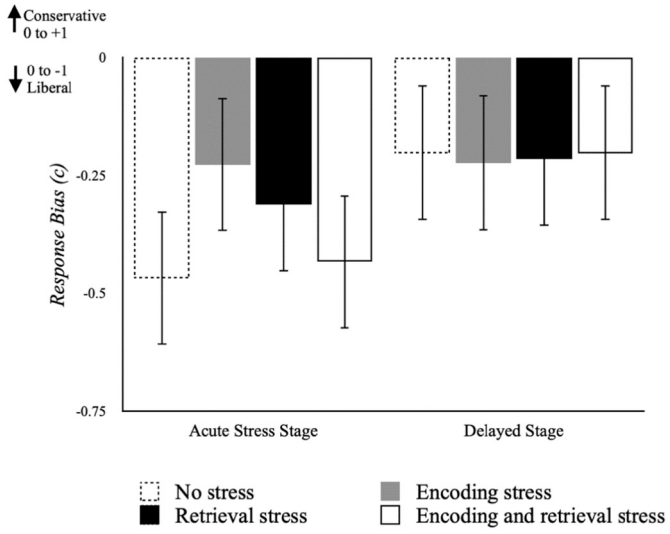


**Fig. 5.** Response bias (c) across stress conditions for acute stress and delayed stages in Experiment 1.
*Note.* Error bars = 95% confidence intervals.

that participants see different images of the same faces at encoding and retrieval (as recommended, e.g., Burton, 2013), rather than seeing the exact same stimulus in both phases (e.g., as with words or images). Moreover, whereas participants saw each face twice during encoding in Experiment 1, they only encoded each face once in Experiment 2. Consequently, participants showed higher hit rates for faces in Experiment 1 than in Experiment 2. Although overall accuracy and sensitivity rates in Experiment 2 do not suggest floor effects for face recognition, the hit rates for faces in Experiment 2 were around chance level. Thus, although these hit rates are comparable to similar work with much shorter retention intervals (Davis et al., 2019; Pezdek et al., 2020), it is possible that floor effects for face recognition may not have allowed us to fully differentiate between individuals performing at such low levels for hit rate. Even so, results from Experiment 1, where hit rates did not show floor effects, similarly suggested an absence of acute stress on face recognition performance.

Other limitations concern factors that could be further assessed in future research include sex/sex hormone status and stimuli valance, as past research suggest differences in stress responses depending on sex hormone status (e.g., Cahill, 2012; Hidalgo, Pulopulos, & Salvador, 2019; Kudielka et al., 2009; Nielsen, Segal, Worden, Yim, & Cahill, 2013). Although our sample does not have enough statistical power to directly examine gender differences in a meaningful way, other samples may show reliable differences in cortisol responses or measures of memory performance depending on sex hormone status (cf. Shields, 2020). Additionally, future research could examine possible interactions between stimulus and participant gender (e.g., Herlitz & Lovén, 2013). Furthermore, a more thorough examination of stimulus valence could advance this area of experimentation. Although meta-analyses found no evidence that stimulus valence moderates the effects of acute encoding stress on memory performance (Shields et al., 2017, but see Schwabe, Bohringer, et al., 2008; Zoladz et al., 2011), larger stress effects have emerged when examining retrieval stress effects on memory performance (e.g., Kuhlmann et al., 2005; Schönfeld et al., 2014). Examining the full range of valence options (e.g., neutral, positive, negative) and directly comparing faces and words would allow for a firmer conclusion. Finally, it is possible that smaller effects of acute stress on face recognition performance are present, but could not be detected in our analyses, which had power to detect medium-large (Experiment 1) or medium (Experiment 2) effect sizes.

A final limitation concerns generalizability to other populations. Our sample consisted exclusively of white participants and white target and filler faces, to limit any effects of the own-race bias (e.g., Meissner & Brigham, 2001). Exploring these research questions in more diverse samples, as well as examining acute stress effects on face recognition in cross-race contexts, is an important future step that has not yet been addressed by empirical research. Additionally, stress may affect children (Deffenbacher et al., 2004) or older adults (Hidalgo et al., 2019; Smith, Dijkstra, Gordon, Romero, & Thomas, 2019) differently to how it affects younger adults, the population examined in our sample. Thus, the current findings may not apply to other age groups.

Crucially, future work on acute stress and face recognition memory should examine the encoding and retrieval memory phases, both fundamentally and in more applied settings. Understanding acute stress effects during both memory phases is valuable for applied forensic settings. To do so, ensuring that suitable manipulation checks and retention intervals are in place is essential. In addition, further exploring the intricacies of the acute stress timeline (e.g., stress stage) will be important for refining our knowledge of the specific situations when acute stress may enhance, impair, or not affect face recognition performance. Finally, careful use of terminology is critical, particularly when researchers use complex scenarios where several factors are likely at play, including stress, attention, cognitive load, and other factors (e.g., Wulff & Thomas, 2021). Disentangling these potential cumulative effects will help provide a much clearer picture on this applied topic.

### 4.1. Conclusion

To conclude, our findings add to a growing body of research demonstrating an absence of stress effects on recognition performance. That is, previous eyewitness and fundamental experiments alike have obtained null effects of acute stress at encoding for face identification tasks (Sauerland et al., 2016), pictures (e.g., Goldfarb, Tompary, Davachi, & Phelps, 2019), and words (e.g., Domes et al., 2004). Additionally, our results regarding retrieval stress support other findings that retrieval stress did not affect face recognition performance (Li et al., 2014). Presently, researchers are beginning to use more robust methodology to investigate stress and cognition (see Shields, 2020), and aligning with best practice methods for hypothesis testing, such as conducting informed power analyses to gather sufficient sample sizes and following preregistered analysis plans. These scientific improvements will make for a stronger evidence base in this complex literature, as reduced power can lead to overestimations in statistically significant differences. A corpus of well-powered, reliable experiments will allow for a fuller understanding of the stress-memory relationship in general and specifically when examining memory for faces.

## CRediT authorship contribution statement

**Carey Marr:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Visualization, Project administration, Funding acquisition. **Conny W.E.M. Quaedflieg:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Henry Otgaar:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Lorraine Hope:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Melanie Sauerland:** Conceptualization, Methodology, Writing – review & editing, Supervision.

## Declaration of competing interest

We have no known conflict of interest to disclose.

*Acknowledgements*

This work was supported by a fellowship awarded from the Erasmus Mundus Joint Doctorate Program The House of Legal Psychology (EMJD-LP) with Framework Partnership Agreement (FPA) 2013-0036 and Specific Grant Agreement (SGA) 532473-EM-5-2017-1-NL-ERA MUNDUS-EMJD to Carey Marr, and partly supported with an award from the American Psychology-Law Society Student Grants-in-Aid to Carey Marr.

*Data statement*

Anonymized data and supplementary materials have been made publicly available on OSF and can be accessed with the following links:

Experiment 1: https://osf.io/xe3wy/?view_only=2f40cea79895475a9b0d1f8a551d18f5

Experiment 2: https://osf.io/m2jc7/?view_only=d74c452bd03641638785becf0091d4e6

The study materials have not been publicly shared because we do not have ethical approval to share the pictures of the face recognition database volunteers. However, all materials are described in depth in the manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.actpsy.2021.103376.

## References

Abercrombie, H. C., Kalin, N. H., Thurow, M. E., Rosenkranz, M. A., & Davidson, R. J. (2003). Cortisol variation in humans affects memory for emotionally laden and neutral information. *Behavioral Neuroscience, 117*(3), 505–516. https://doi.org/10.1037/0735-7044.117.3.505

Bailis, K. L., & Mueller, J. H. (1981). Anxiety, feedback, and self-reference in face recognition. *Motivation and Emotion, 5*(1), 85–96. https://doi.org/10.1007/bf00993664

Bornstein, B. H., Hullman, G., & Miller, M. K. (2013). Stress, trauma, and wellbeing in the legal system: Where do we go from here? In M. K. Miller, & B. H. Bornstein (Eds.), *Stress, trauma, and wellbeing in the legal system* (pp. 293–309). Oxford Press. https://doi.org/10.1093/acprof:oso/9780199829996.001.0001.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical report C-1*. The Center for Research in Psychophysiology, University of Florida.

Brigham, J. C., Maass, A., Martinez, D., & Wittenberger, G. (1983). The effect of arousal on facial recognition. *Basic and Applied Social Psychology, 4*(3), 279–293. https://doi.org/10.1207/s15324834basp0403_6

Buchanan, T. W., & Tranel, D. (2008). Stress and emotional memory retrieval: Effects of sex and cortisol response. *Neurobiology of Learning and Memory, 89*(2), 134–141. https://doi.org/10.1016/j.nlm.2007.07.003

Buchanan, T. W., Tranel, D., & Adolphs, R. (2006). Impaired memory retrieval correlates with individual differences in cortisol response but not autonomic response. *Learning & Memory, 13*(3), 382–387. https://doi.org/10.1101/lm.206306

Buckhout, R., Alper, A., Chern, S., Silverberg, G., & Slomovits, M. (1974). Determinants of eyewitness performance on a lineup. *Bulletin of the Psychonomic Society, 4*(3), 191–192. https://doi.org/10.3758/bf03334241

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology, 66*(8), 1467–1485. https://doi.org/10.1080/17470218.2013.800125

Cahill, L. (2012). A half-truth is a whole lie: On the necessity of investigating sex influences on the brain. *Endocrinology, 153*(6), 2541–2543. https://doi.org/10.1210/en.2011-2167

Cahill, L., Gorski, L., & Le, K. (2003). Enhanced human memory consolidation with post-learning stress: Interaction with the degree of arousal at encoding. *Learning & Memory, 10*(4), 270–274. https://doi.org/10.1101/lm.62403

Clifford, B. R., & Hollin, C. R. (1981). Effects of the type of incident and the number of perpetrators on eyewitness memory. *Journal of Applied Psychology, 66*(3), 364–370. https://doi.org/10.1037/0021-9010.66.3.364

Coleman, A., Freeman, P., Steel, S., & Shennan, A. (2006). Validation of the Omron 705IT (HEM-759-E) oscillometric blood pressure monitoring device according to the British Hypertension Society protocol. *Blood Pressure Monitoring, 11*(1), 27–32. https://doi.org/10.1097/01.mbp.0000189788.05736.5f

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). The reliability of eyewitness identification: The role of system and estimator variables. *Law and Human Behavior, 11*(3), 233–258. https://doi.org/10.2307/1393634

Davis, S. D., Peterson, D. J., Wissman, K. T., & Slater, W. A. (2019). Physiological stress and face recognition: Differential effects of stress on accuracy and the confidence-accuracy relationship. *Journal of Applied Research in Memory and Cognition, 8*, 367–375. https://doi.org/10.1016/j.jarmac.2019.05.006

de Quervain, D. J.-F., Henke, K., Aerni, A., Treyer, V., McGaugh, J. L., Berthold, T., Nitsch, R. M., Buck, A., Roozendaal, B., & Hock, C. (2003). Glucocorticoid-induced impairment of declarative memory retrieval is associated with reduced block flow in the medial temporal lobe. *European Journal of Neuroscience, 17*(6), 1296–1302. https://doi.org/10.1046/j.1460-9568.2003.02542.x

de Quervain, D. J.-F., Roozendaal, B., & McGaugh, J. L. (1998). Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature, 394*(6695), 787–790. https://doi.org/10.1038/29542

de Quervain, D. J.-F., Roozendaal, B., Nitsch, R. M., McGaugh, J. L., & Hock, C. (2000). Acute cortisone administration impairs retrieval of long-term declarative memory in humans. *Nature Neuroscience, 3*(4), 313–314. https://doi.org/10.1038/73873

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28*(6), 687–706. https://doi.org/10.1007/s10979-004-0565-x

Diamond, D. M., Campbell, A. M., Park, C. R., Halonen, J., & Zoladz, P. R. (2007). The temporal dynamics model of emotional processing: A synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the Yerkes-Dodson Law. *Neural Plasticity, 2007*, Article 60803. https://doi.org/10.1155/2007/60803

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*(2), 107–117. https://doi.org/10.1037/0096-3445.115.2.107

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Domes, G., Heinrichs, M., Reichwald, U., & Hautzinger, M. (2002). Hypothalamic-pituitary-adrenal axis reactivity to psychological stress and memory in middle-aged women: High responders exhibit enhanced declarative memory performance. *Psychoneuroendocrinology, 27*(7), 843–853. https://doi.org/10.1016/s0306-4530(01)00085-3

Domes, G., Heinrichs, M., Rimmele, U., Reichwald, U., & Hautzinger, M. (2004). Acute stress impairs recognition for positive words—Association with stress-induced cortisol secretion. *Stress, 7*(3), 173–181. https://doi.org/10.1080/10253890412331273213

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/bf03193146

Gagnon, S. A., & Wagner, A. D. (2016). Acute stress and episodic memory: Neurobiological mechanisms and behavioral consequences. *Annals of the New York Academy of Sciences, 1369*(1), 55–75. https://doi.org/10.1111/nyas.12996

Goldfarb, E. V., Tompary, A., Davachi, L., & Phelps, E. A. (2019). Acute stress throughout the memory cycle: Diverging effects on associative and item memory. *Journal of Experimental Psychology: General, 148*(1), 13–29. https://doi.org/10.1037/xge0000472

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4*(9), 330–337. https://doi.org/10.1016/s1364-6613(00)01519-9

Hellhammer, J., & Schubert, M. (2012). The physiological response to Trier Social Stress Test relates to subjective measures of stress during but not before or after the test. *Psychoneuroendocrinology, 37*(1), 119–124. https://doi.org/10.1016/jpsyneuen.2011.05.012

Henckens, M. J. A. G., Hermans, E. J., Pu, Z., Joëls, M., & Fernández, G. (2009). Stressed memories: How acute stress affects memory formation in humans. *The Journal of Neuroscience, 29*(32), 10111–10119. https://doi.org/10.1523/JNEUROSCI.1184-09.2009

Herlitz, A., & Lovén, J. (2013). Sex differences and the own-gender bias in face recognition: A meta-analytic review. *Visual Cognition, 21*(9–10), 1306–1336. https://doi.org/10.1080/13506285.2013.823140

Hermans, E. J., Henckens, M. J., Joëls, M., & Fernández, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends in Neurosciences, 37*(6), 304–314. https://doi.org/10.1016/j.tins.2014.03.006

Het, S., Ramlow, G., & Wolf, O. T. (2005). A meta-analytic review of the effects of acute cortisol administration on human memory. *Psychoneuroendocrinology, 30*(8), 771–784. https://doi.org/10.1016/j.psyneuen.2005.03.005

Hidalgo, V., Pulopulos, M. M., & Salvador, A. (2019). Acute psychosocial stress effects on memory performance: Relevance of age and sex. *Neurobiology of Learning and Memory, 157*, 48–60. https://doi.org/10.1016/j.nlm.2018.11.013

Hope, L., Blocksidge, D., Gabbert, F., Sauer, J. D., Lewinski, W., Mirashi, A., & Atuk, E. (2016). Memory and the operational witness: Police officer recall of firearms encounters as a function of active response role. *Law and Human Behavior, 40*(1), 23–35. https://doi.org/10.1037/lhb0000159

Hoscheidt, S. M., LaBar, K., Ryan, L., Jacobs, W. J., & Nadel, L. (2014). Encoding negative events under stress: High subjective arousal is related to accurate emotional memory despite misinformation exposure. *Neurobiology of Learning and Memory, 112*, 237–247. https://doi.org/10.1016/j.nlm.2013.09.008

Hulse, L. M., & Memon, A. (2006). Fatal impact? The effects of emotional arousal and weapon presence on police officers' memories for a simulated crime. *Legal and Criminological Psychology, 11*(2), 313–325. https://doi.org/10.1348/135532505X58062

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving, 7*(1), 2–9. https://doi.org/10.7771/1932-6246.1167

JASP Team. (2020). *JASP (version 0.13.1) [computer software]*.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.

Joëls, M., & Baram, T. Z. (2009). The neuro-symphony of stress. *Nature Reviews Neuroscience, 10*(6), 459–466. https://doi.org/10.1038/nrn2632

Joëls, M., Fernández, G., & Roozendaal, B. (2011). Stress and emotional memory: A matter of timing. *Trends in Cognitive Sciences, 15*(6), 280–288. https://doi.org/10.1016/j.tics.2011.04.004

Joëls, M., Pu, Z., Wiegert, O., Oitzl, M. S., & Krugers, H. J. (2006). Learning under stress: How does it work? *Trends in Cognitive Sciences, 10*(4), 152–158. https://doi.org/10.1016/j.tics.2006.02.002

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences, 361*(1476), 2109–2128. https://doi.org/10.1098/rstb.2006.1934

Kirschbaum, C., Kudielka, B. M., Gaab, J., Schommer, N. C., & Hellhammer, D. H. (1999). Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic Medicine, 61*(2), 154–162. https://doi.org/10.1097/00006842-199903000-00006

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The "Trier Social Stress Test": A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*(1–2), 76–81. https://doi.org/10.1159/000119004

Kramer, T. H., Buckhout, R., & Eugenio, P. (1990). Weapon focus, arousal, and eyewitness memory: Attention must be paid. *Law and Human Behavior, 14*(2), 167–184. https://doi.org/10.2307/1393597

Kudiekla, B. M., Hellhammer, D. H., & Wust, S. (2009). Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology, 34*, 2–18. https://doi.org/10.1016/j.psyneuen.2008.10.004

Kuhlmann, S., Piel, M., & Wolf, O. T. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *The Journal of Neuroscience, 25*(11), 2977–2982. https://doi.org/10.1523/jneurosci.5139-04-2005

Li, S., Weerda, R., Guenzel, F., Wolf, O. T., & Thiel, C. M. (2013). ADRA2B genotype modulates effects of acute psychosocial stress on emotional memory retrieval in healthy young men. *Neurobiology of Learning and Memory, 103*, 11–18. https://doi.org/10.1016/j.nlm.2013.03.006

Li, S., Weerda, R., Milde, C., Wolf, O. T., & Thiel, C. M. (2014). Effects of acute psychosocial stress on neural activity to emotional and neutral faces in a face recognition paradigm. *Brain Imaging and Behavior, 8*(4), 598–610. https://doi.org/10.1007/s11682-013-9287-3

Maass, A., & Kohnken, G. (1989). Eyewitness identification: Simulating the "weapon effect.". *Law and Human Behavior, 13*(4), 397–408. https://doi.org/10.1007/bf01056411

Marr, C., Otgaar, H., Sauerland, M., Quaedflieg, C. W. E. M., & Hope, L. (2020). The effects of stress on eyewitness memory: A survey of memory experts and laypeople. *Memory & Cognition, 49*(3), 401–421. https://doi.org/10.3758/s13421-020-01115-4

McGaugh, J. L. (2015). Consolidating memories. *Annual Review of Psychology, 66*, 1–24. https://doi.org/10.1146/annurev-psych-010814-014954

McGugin, R. W., Newton, A. T., Gore, J. C., & Gauthier, I. (2014). Robust expertise effects in right FFA. *Neuropsychologia, 63*, 135–144. https://doi.org/10.1016/j.neuropsychologia.2014.08.02

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in face memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*(1), 3–35 (doi:1037/1076-8971.7.1.3).

Miller, R., Plessow, F., Kirschbaum, C., & Stalder, T. (2013). Classification criteria for distinguishing cortisol responders from nonresponders to psychosocial stress: Evaluation of salivary cortisol pulse detection in panel designs. *Psychosomatic Medicine, 75*(9), 832–840. https://doi.org/10.1097/PSY.0000000000000002

Morgan, C. A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baranoski, M., & Southwick, S. M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry, 27*(3), 265–279. https://doi.org/10.1016/j.ijlp.2004.03.004

Mueller, J. H., Bailis, K. L., & Goldstein, A. G. (1979). Depth of processing and anxiety in facial recognition. *British Journal of Psychology, 70*(4), 511–515. https://doi.org/10.1111/j.2044-8295.1979.tb01724.x

Nielsen, S. E., Segal, S. K., Worden, I. V., Yim, I. S., & Cahill, L. (2013). Hormonal contraception use alters stress responses and emotional memory. *Biological Psychology, 92*(2), 257–266 (doi:j.biopsycho.2012.10.007).

Nowicki, S., Winograd, E., & Millard, B. A. (1979). Memory for faces: A social learning analysis. *Journal of Research in Personality, 13*(4), 460–468. https://doi.org/10.1016/0092-6566(79)90008-4

Paul, M., Lech, R. K., Scheil, J., Dierolf, A. M., Suchan, B., & Wolf, O. T. (2016). Acute stress influences the discrimination of complex scenes and complex faces in young healthy men. *Psychoneuroendocrinology, 66*, 125–129. https://doi.org/10.1016/j.psyneuen.2016.01.007

Payne, J. D., Jackson, E. D., Ryan, L., Hoscheidt, S., Jacobs, J. W., & Nadel, L. (2006). The impact of stress on neutral and emotional aspects of memory. *Memory, 14*(1), 1–6. https://doi.org/10.1080/09658210500139176

Peters, D. P. (1988). Eyewitness memory in a natural setting. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (pp. 89–94). Wiley.

Pezdek, K., Abed, E., & Cormia, A. (2020). Elevated stress impairs the accuracy of eyewitness memory but not the confidence-accuracy relationship. *Journal of Experimental Psychology: Applied*. https://doi.org/10.1037/xap000316. Advance online publication.

Quaedflieg, C. W. E. M., Meyer, T., van Ruitenbeek, P., & Smeets, T. (2017). Examining habituation and sensitization across repetitive laboratory stress inductions using the MAST. *Psychoneuroendocrinology, 77*, 175–181. https://doi.org/10.1016/j.psyneuen.2016.12.009

Quaedflieg, C. W. E. M., & Schwabe, L. (2018). Memory dynamics under stress. *Memory, 26*(3), 364–376. https://doi.org/10.1080/09658211.2017.1338299

Quaedflieg, C. W. E. M., Schwabe, L., Meyer, T., & Smeets, T. (2013). Time dependent effects of stress prior to encoding on event-related potentials and 24 h delayed retrieval. *Psychoneuroendocrinology, 38*(12), 3057–3069. https://doi.org/10.1016/j.psyneuen.2013.09.002

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111–163. https://doi.org/10.2307/271063

Sato, W., & Yoshikawa, S. (2013). Recognition memory for faces and scenes. *The Journal of General Psychology, 140*(1), 1–15. https://doi.org/10.1080/00221309.2012.710275

Sauerland, M., Raymaekers, L. H. C., Otgaar, H., Memon, A., Waltjen, T. T., Nivo, M., … Smeets, T. (2016). Stress, stress-induced cortisol responses, and eyewitness identification performance. *Behavioral Sciences & the Law, 34*(4), 580–594. https://doi.org/10.1002/bsl.2249

Schönfeld, P., Ackermann, K., & Schwabe, L. (2014). Remembering under stress: Different roles of autonomic arousal and glucocorticoids in memory retrieval. *Psychoneuroendocrinology, 39*(1), 249–256. https://doi.org/10.1016/j.psychneuen.2013.09.020

Schwabe, L., Bohringer, A., Chatterjee, M., & Schachinger, H. (2008). Effects of pre-learning stress on memory for neutral, positive, and negative words: Differential roles of cortisol and autonomic arousal. *Neurobiology of Learning and Memory, 90*(1), 44–53. https://doi.org/10.1016/j.nlm.2008/02.002

Schwabe, L., Haddad, L., & Schachinger, H. (2008). HPA axis activation by a socially evaluated cold-pressor test. *Psychoneuroendocrinology, 33*(6), 890–895. https://doi.org/10.1016/j.psyneuen.2008.03.001

Schwabe, L., Joëls, M., Roozendaal, B., Wolf, O. T., & Oitzl, M. S. (2012). Stress effects on memory: An update and integration. *Neuroscience & Biobehavioral Reviews, 36*(7), 1740–1749. https://doi.org/10.1016/j.neubiorev.2011.07.002

Schwabe, L., & Wolf, O. T. (2010). Learning under stress impairs memory formation. *Neurobiology of Learning and Memory, 93*(2), 183–188. https://doi.org/10.1016/j.nlm.2009.09.009

Schwabe, L., & Wolf, O. T. (2014). Timing matters: Temporal dynamics of stress effects on memory retrieval. *Cognitive, Affective, & Behavioral Neuroscience, 14*(3), 1041–1048. https://doi.org/10.3758/s13415-014-0256-0

Shermohammed, M., Davidow, J. Y., Somerville, L. H., & Murty, V. P. (2019). Stress impacts the fidelity but not strength of emotional memories. *Brain and Cognition, 133*, 33–41. https://doi.org/10.1016/j.bandc.2018.09.001

Shields, G. S. (2020). Stress and cognition: A user's guide to designing and interpreting studies. *Psychoneuroendocrinology, 112*, Article 104475. https://doi.org/10.1016/j.psyneuen.2019.104475

Shields, G. S., Sazma, M. A., McCullough, A. M., & Yonelinas, A. P. (2017). The effects of acute stress on episodic memory: A meta-analysis and integrative review. *Psychological Bulletin, 143*(6), 636–675. https://doi.org/10.1037/bul0000100

Smeets, T., Cornelisse, S., Quaedflieg, C. W. E. M., Meyer, T., Jelicic, M., & Merckelbach, H. (2012). Introducing the Maastricht Acute Stress Test (MAST): A quick and non-invasive approach to elicit robust autonomic and glucocorticoid stress responses. *Psychoneuroendocrinology, 37*(12), 1998–2008. https://doi.org/10.1016/j.psyneuen.2012.04.012

Smeets, T., Otgaar, H., Candel, I., & Wolf, O. T. (2008). True or false? Memory is differentially affected by stress-induced cortisol elevations and sympathetic activity at consolidation and retrieval. *Psychoneuroendocrinology, 33*(10), 1378–1386. https://doi.org/10.1016/j.psyneuen.2008.07.009

Smith, A. M., Dijkstra, K., Gordon, L. T., Romero, M., & Thomas, A. K. (2019). An investigation into the impact of acute stress on encoding in older adults. *Aging, Neuropsychology, and Cognition, 26*(5), 749–766. https://doi.org/10.1080/13825585.2018.1524438

Strahler, J., Skoluda, N., Kappert, M. B., & Nater, U. M. (2017). Simultaneous measurement of salivary cortisol and alpha-amylase: Application and recommendation. *Neuroscience & Biobehavioral Reviews, 83*, 657–677. https://doi.org/10.1016/j.neubiorev.2017.08.015

Tooley, V., Brigham, J. C., Maas, A., & Bothwell, R. K. (1987). Facial recognition: Weapon effect and attentional focus. *Journal of Applied Social Psychology, 17*(10), 845–859. https://doi.org/10.1111/j.1559-1816.1987.tb00294.x

Valentine, T., & Mesout, J. (2008). Eyewitness identification under stress in the London Dungeon. *Applied Cognitive Psychology, 23*(2), 151–161. https://doi.org/10.1002/acp.1463

Vogel, S., & Schwabe, L. (2016). Stress in the zoo: Tracking the impact of stress on memory formation over time. *Psychoneuroendocrinology, 71*, 64–72. https://doi.org/10.1016/j.psyneuen.2016.04.027

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063. https://doi.org/10.1037/0022-3514.54.6.1063

Wiemers, U. S., Sauvage, M. M., Schoofs, D., Hamacher-Dang, T. C., & Wolf, O. T. (2013). What we remember from a stressful episode. *Psychoneuroendocrinology, 38*(10), 2268–2277. https://doi.org/10.1016/j.psyneuen.2013.04.015

Wolf, O. T. (2012). Immediate recall influences the effects of pre-encoding stress on emotional episodic long-term memory consolidation in healthy young men. *Stress, 15*(3), 272–280. https://doi.org/10.3109/10253890.2011.622012

Wolf, O. T. (2017). Stress and memory retrieval: Mechanisms and consequences. *Current Opinion in Behavioral Sciences, 14*, 40–46. https://doi.org/10.1016/cobeha.2016.12.001

Woodhead, M. M., & Baddeley, A. D. (1981). Individual differences and memory for faces, pictures, and words. *Memory & Cognition, 9*(4), 368–370. https://doi.org/10.3758/bf03197561

Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology, 71*, 291–301. https://doi.org/10.1037/0021-9010.71.2.291

Zoladz, P. R., Clark, B., Warnecke, A., Smith, L., Tabar, J., & Talbot, J. N. (2011). Pre-learning stress differentially affects long-term memory for emotional words, depending on temporal proximity to the learning experience. *Physiology & Behavior, 103*(5), 467–476. https://doi.org/10.1016/j.physbeh.2011.01.016

Zoladz, P. R., Kalchik, A. E., Hoffman, M. M., Aufdenkampe, R. L., Lyle, S. M., Peters, D. M., Brown, C. M., Cadle, C. E., Scharf, A. R., Dailey, A. M., Wolters, N. E., Talbot, J. N., & Rorabaugh, B. R. (2014). ADRA2B deletion variant selectively predicts stress-induced enhancement of long-term memory in females. *Psychoneuroendocrinology, 48*, 111–122. https://doi.org/10.1016/j.psyneun.2014.06.01