



Mitigating the negative effects of retrieval stress on memory: an arousal reappraisal intervention

Carey Marr, Melanie Sauerland, Henry Otgaar, Conny W.E.M. Quaedflieg & Lorraine Hope

To cite this article: Carey Marr, Melanie Sauerland, Henry Otgaar, Conny W.E.M. Quaedflieg & Lorraine Hope (2021): Mitigating the negative effects of retrieval stress on memory: an arousal reappraisal intervention, *Memory*, DOI: [10.1080/09658211.2021.1893750](https://doi.org/10.1080/09658211.2021.1893750)

To link to this article: <https://doi.org/10.1080/09658211.2021.1893750>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 09 Mar 2021.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Mitigating the negative effects of retrieval stress on memory: an arousal reappraisal intervention

Carey Marr ^{a,b}, Melanie Sauerland ^a, Henry Otgaar ^{a,c}, Conny W.E.M. Quaedflieg ^a and Lorraine Hope ^b

^aFaculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands; ^bDepartment of Psychology, University of Portsmouth, Portsmouth, UK; ^cFaculty of Law, Catholic University of Leuven, Leuven, Belgium

ABSTRACT

In a preregistered experiment, we examined the efficacy of arousal reappraisal as an intervention for reducing the negative effects of stress at retrieval on memory. Participants ($N = 177$) were semi-randomly assigned to one of three conditions: a Stress-intervention condition, a Stress-placebo condition, and a No-stress-placebo control condition. Participants viewed four images of complex, mildly negatively valenced scenes. One day later, they received an arousal reappraisal intervention or placebo before exposure to a laboratory stressor (or a control version for the No-stress condition). Participants were then tested on their memory of the images using a free recall instruction and multiple-choice recognition questions. As expected, negative affect and blood pressure increased for the stress conditions but not the control condition. Contrary to our hypotheses, memory performance did not statistically significantly differ between the Stress-placebo condition and the No-stress-placebo control condition, indicating a lack of negative effects of acute retrieval stress on memory. Furthermore, we also found no statistically significant differences between the Stress-intervention condition and Stress-placebo condition in terms of memory performance, suggesting that the intervention did not assist with enhancing memory. We integrate interpretations of the findings from this study with a discussion of avenues for future research in this area.

ARTICLE HISTORY

Received 10 September 2018
Accepted 12 February 2021

KEYWORDS


stress; memory; eyewitness; arousal reappraisal; free recall; recognition; memory enhancement

When a person experiences acute stress, their autonomic system is activated, leading to the release of adrenaline and symptoms such as increased heart rate, blood pressure, and sweating. In addition to this autonomic arousal, stress also activates the hypothalamic–pituitary–adrenal (HPA) axis. This activation leads to the release of cortisol, which consists of both genomic and non-genomic effects (Joëls et al., 2011; Quaedflieg & Schwabe, 2018). The non-genomic effects influence the basolateral amygdala (BLA) about 15–20 min after an acute stressor is experienced and continue for around one hour. The genomic effects are more delayed, beginning around 60–90 min post-stressor and lasting hours or even days. These cortisol effects typically impair memory retrieval (de Quervain et al., 1998; Shields et al., 2017; Wolf, 2017), and these negative effects of stress on retrieval are most often found during the cortisol peak (around 20 min post-stressor; e.g., Schwabe & Wolf, 2014). This is the case for both recall (Kuhlmann et al., 2005; Schönfeld et al., 2014; Smeets et al., 2008) and recognition tasks (Li et al., 2013; Schwabe & Wolf,

2014), although research suggests stronger effects for recall than recognition (see Gagnon & Wagner, 2016). Research has investigated memory for different kinds of stimuli including words (Kuhlmann et al., 2005; Schönfeld et al., 2014; Schwabe & Wolf, 2014; Smeets et al., 2008), pictures (Schönfeld et al., 2014), and faces (Li et al., 2013). Generally, larger effects are often found for emotional memory compared to neutral memory (e.g., Kuhlmann et al., 2005; Schönfeld et al., 2014).

The effect of acute stress before retrieval could be problematic during police interviews with witnesses. Drawing on past research concerning how acute stress negatively impacts memory retrieval (e.g., Shields et al., 2017), it is likely that witnesses who experience heightened pressure during an interview may show reduced quantity and accuracy of information. This could happen when an interview is conducted in a harsh, aggressive and suggestive manner. However, even if police are sensitive and do not explicitly set out to impose stress on witnesses during interviews, witnesses may still find the situation stressful. For example, it is

CONTACT Carey Marr  carey.marr@maastrichtuniversity.nl  Universiteitssingel 40, Maastricht 6229 ER, the Netherlands

 Supplemental data for this article can be accessed <https://doi.org/10.1080/09658211.2021.1893750>

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

possible that witnesses who are uncertain about how the interview process will transpire (Sydeman et al., 1997), believe they could be implicated in the crime, are fearful of the perpetrators they are reporting about, or suffer from social or evaluation anxiety (particularly when dealing with authority figures; e.g., Leitenberg, 1990) may experience increased stress levels during a police interview. Although no research has yet examined this issue in the context of police interview settings, research examining other interview settings suggests increased levels of stress for those in the interviewee position (e.g., McCarthy & Goffin, 2004; Posthuma et al., 2002).

Recently, researchers have detected ways to reduce the negative effects of stress on memory retrieval. For example, Smith et al. (2016) found that participants who engaged in retrieval practice (i.e., taking practice tests) were immune to the negative effects of stress, while those who simply restudied the materials showed the typical memory impairments. Because stress levels cannot always be avoided or reduced, one tactic for managing stress is to reframe the way individuals view arousal. Research on cognitive reappraisal suggests that altering the way one interprets situations can change subsequent emotional or even physiological reactions (e.g., Lazarus, 1991). More recently, researchers have started to apply such cognitive reappraisal strategies to stressful situations. One reappraisal strategy, *arousal reappraisal*, has been developed from the theoretical biopsychosocial model of challenge and threat (Blascovich & Mendes, 2010; Blascovich & Tomaka, 1996). This model suggests that in motivated performance situations, arousal can lead individuals to either a challenge or a threat state, which exist on two opposite ends of a spectrum. In a threat state, individuals believe they do not have appropriate resources to meet the situational demands of the stressful situation. This state leads to avoidance of the situation, increased negative affect, autonomic arousal, and HPA axis activation, and impaired cognitive performance (e.g., Jamieson et al., 2018; Seery, 2011). In a challenge state, individuals believe they are able to cope with the situational demands, leading them to approach the situation. Individuals in this state report more excitement and increased self-esteem, physiologically display heightened autonomic arousal but moderate HPA activation, and show enhanced cognitive performance (e.g., Jamieson et al., 2018; Mendes & Park, 2014). Therefore, arousal reappraisal does not decrease stress reactions, encourage relaxation, or reduce sympathetic arousal, but rather focuses on shifting the experience of stress from a threat state to a challenge state (Jamieson et al., 2018). Arousal reappraisal, consisting of advice given to individuals before they encounter stress, aims to guide individuals towards a challenge state by helping them view their feelings and bodily response to stress (e.g., sweaty palms, increased heart rate, etc.) as adaptive and useful rather than harmful.

Past research shows that arousal reappraisal helps participants perform better during objective laboratory stress tests compared with counterparts who do not receive arousal reappraisal instructions (e.g., giving better quality speeches during the Trier Social Stress Test; Beltzer et al., 2014). Researchers have also found similar beneficial effects in more applied stressful settings, such as on test-taking performance (e.g., performing better on standardised math examinations; Jamieson et al., 2010). These findings coupled with research showing that arousal reappraisal elicits more moderate cortisol reactions (e.g., Blascovich, 2008; Mendes & Park, 2014) indicate that arousal reappraisal could potentially reduce the negative effects of stress before retrieval on memory. However, no research has directly investigated this possibility. Additionally, research suggests that arousal reappraisal may have other positive benefits, including reduction of negative affect and anxiety, and increasing coping ability (Jamieson et al., 2018). Though not related to memory performance, improving affect and coping ability – for example for witnesses who have just experienced a crime and a subsequent police interview – would be an added benefit of such an intervention.

The current experiment applied an arousal reappraisal intervention before a stressful situation to test the hypothesis that such an intervention can have a positive effect on recall and recognition performance. To test this research question, we included three conditions: (1) a Stress-intervention condition, in which participants received the arousal reappraisal intervention and were exposed to a stressor prior to the memory test, (2) a Stress-placebo condition, in which participants received the placebo and were exposed to a stressor prior to the memory test, and (3) a No-stress-placebo control condition, in which participants received the placebo and were not stressed prior to the memory test.

Hypotheses

- (1) We predicted that participants in the No-stress-placebo control condition would report more details in total, more correct details, but fewer incorrect details, and hence display better recall and recognition accuracy (proportion of correct details), and display better sensitivity (greater d' scores) than participants in the Stress-placebo condition.
- (2) We predicted that participants in the Stress-intervention condition would report more details in total, more correct details, but fewer incorrect details, and hence display better recall and recognition accuracy, and display better sensitivity than participants in the Stress-placebo condition. Thus, we predicted that the negative effect of retrieval stress on memory (i.e., reflected in the reporting of fewer details and less accurate information) would be reduced for those who received the intervention.

Method

This study, including its method and analyses, was accepted in principle as a registered report. The accepted Stage 1 proposal, as well as all study materials and data, can be found at https://osf.io/8a9sq/?view_only=ce42932ad0cb415f9934b64b1d7928bf.

Participants

Based on a priori power analyses conducted using G*Power (Faul et al., 2007), a one-way fixed-effects ANOVA with 90% power, $\alpha = .05$, and a medium effect size of $f = .27$ (based on past relevant work including Domes et al., 2004; Hidalgo et al., 2015; Kuhlmann et al., 2005; Schwabe & Wolf, 2009, 2014, Experiment 2; Smeets, 2011; Tollenaar et al., 2009), the target sample size was $N = 177$. 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 (e.g., Quaedflieg et al., 2013; Shields, 2020; Strahler et al., 2017), we screened for and excluded participants who habitually smoked (>15 cigarettes per day), drank alcohol (>15 drinks per week), or used drugs (more than once per week). Participants were also excluded for a variety of other health reasons (i.e., BMI < 17 and >30, use of medications containing cortisone, recent vaccinations, psychological treatments, cardiovascular problems, or endocrine disorders). Because sex hormones can affect cortisol reactivity (e.g., Kirschbaum et al., 1999; Strahler et al., 2017), females using hormonal contraceptives, females not using hormonal contraceptives, and males were balanced across experimental conditions. Participants' native languages were Dutch, German, or English, and they completed the memory test part of the study in their native language.

In total, 184 participants were tested. Seven participants were not included in the final sample: one because of experimental error, three because they cancelled Session 2, and three because they withdrew from participation during Session 2 (all females), noting that the task was too painful. Data collection continued until we obtained the final sample consisted of 177 participants, as planned. Of note, data were not collected during the COVID-19 lockdown in the Netherlands (March – June 2020), but some data were collected once the COVID-19 pandemic was temporarily improving (July – October 2020). During this time, we continued to semi-randomly distribute participants equally among all three groups, limiting potential COVID-19 related differences between conditions. The final sample included men ($n = 32$) and women ($n = 145$) between the ages of 18 and 34 years ($M = 22.03$, $SD = 3.06$). About half the women used hormonal contraceptives (51.72%) and the other half did not (48.28%). The majority were university students (87.00%), most of whom were completing their undergraduate

degrees (69.50%), with the rest completing a different degree (30.50%).

We tested between 12:00pm and 6:00pm to account for diurnal cortisol levels (Shields, 2020). Participants were asked to follow several rules before Session 2 of the experiment to help control for variability in cortisol levels. 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 two hours prior to the session. Participants received either course credit or 15 euros in gift vouchers as reimbursement. The three participants with the highest memory test scores (combination of greatest quantity of recall details and greatest recall and recognition accuracy) received an additional 15 euros in vouchers upon completion of data collection. This study was approved by the ethical committee of the Faculty of Psychology and Neuroscience at Maastricht University.

Design

We used a one-factorial between-subjects design with three conditions (Stress-intervention, Stress-placebo, and No-stress-placebo control; $n = 59$ in each condition). The dependent variables were recall quantity (total number of details, total number of correct details, and total number of incorrect details), recall accuracy (number of correct details divided by total number of details reported), recognition accuracy (proportion of correct answers), recognition sensitivity (d'), and response bias (c). Participants answered recall questions for two images and recognition questions for two images. The order of recall and recognition questions was counterbalanced.

Materials

Stimuli and retrieval test

Four colour images depicting complex drawings with object- and person-related details were used (Kitchen, Living Room, Traffic, and Bus), displaying mildly negative scenes (e.g., a person falling over, a flat tire, etc.). Based on previous work using these materials (see Sauer & Hope, 2016), each image was shown for 60 s with a 10-second interval. The recall test for two of these images (i.e., Kitchen and Traffic) consisted of an open-ended invitation to report everything they could remember about the images (including people, events, and actions) in as much detail and as accurately as possible. Participants were discouraged from guessing. The recognition test for the other two images (i.e., Living Room and Bus) consisted of 25 questions per image. Each question was presented with five multiple-choice options (e.g., "What colour were the painter's clothes?" (a) Blue, (b) Red, (c) Orange, (d) Green, (e) None of the above). The fifth choice was always "None of the above" and this option was correct 20–25% of the time. Recognition scores were coded as

the proportion of correct answers (i.e., recognition accuracy). Proportions of hits and false alarms were also calculated. Recognition sensitivity (d') was calculated as the difference between the z-scores of proportion of hits and proportion of false alarms, where the larger the d' , the better one's performance. Response bias (c) was calculated as the sum of the z-scores of hits and false alarms divided by two, where 0 represents no bias, values < 0 represent a liberal bias, and values > 0 represent a conservative bias. Each participant answered recall questions for two of the images and recognition questions for the other two images, that is, they gave two free recall accounts and answered 50 recognition questions in total.

Inter-coder reliability

To ensure appropriate levels of inter-coder reliability for free recall, we used a systematic coding protocol and trained two coders for each language (six total). The coding protocol included a list of possible items that may be reported (including person details, objects, actions, etc.) and coders coded against this list, updating and adding details if necessary. Two coders coded 34% of the free recall responses. Intraclass coefficient correlations (ICCs) were calculated based on a single rating, absolute-agreement, two-way random effects model. ICCs ranged between .680 and .938, indicating moderate to excellent reliability (Koo & Li, 2016). Table A in supplementary materials shows the full ICCs analyses.

Maastricht Acute Stress Test

Participants in the stress conditions were exposed to the Maastricht Acute Stress Test (MAST), a validated laboratory stressor that combines blocks of hand immersion into ice-cold water (four degrees Celsius) with socially-evaluated mental arithmetic in front of a critical experimenter who gives negative feedback throughout the task (see Smeets et al., 2012). Additionally, as part of the MAST, participants were told they were being video recorded for later facial expression analysis. Participants in the No-stress-placebo control condition were exposed to the control version of this task, involving hand immersion into room-temperature water (35 degrees Celsius) and basic counting, with no mention of video recording.

Positive and Negative Affect Schedule (PANAS)

The Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) consists of two affect scales that each contain 10 items to measure positive and negative affect. Participants indicate to what extent they feel a certain mood (e.g., interested, excited, nervous, distressed, etc.) at the present moment on a 5-point Likert scale (from *very slightly or not at all* to *extremely*). We used the PANAS to measure positive and negative affect at five timepoints: during Session 1, at the start of Session 2, after the intervention (or placebo), after the (control) MAST, and after the memory test.

Blood pressure

To examine autonomic nervous system activation, systolic and diastolic blood pressure were collected using an Omron Blood Pressure Monitor 705IT (for validation see Coleman et al., 2006). Blood pressure was measured ten times throughout the procedure: during Session 1, baseline during Session 2, before the intervention (or placebo), during the intervention (or placebo), right after the intervention (or placebo), before the (control) MAST, during the (control) MAST, right after the (control) MAST, before the memory test, and during the memory test.

Arousal reappraisal and placebo interventions

Arousal reappraisal materials were adapted from past work (Jamieson et al., 2016, 2018). These materials included summaries of three scientific articles pertaining to what happens when we experience stress. In the arousal reappraisal intervention, these articles emphasised the adaptive benefits of stress and suggested that perceiving such arousal as adaptive would enhance performance. In the placebo version, these articles suggested that ignoring the symptoms of arousal would help performance.

Stress appraisal

A validated measure to evaluate stress appraisals (Mendes et al., 2007) was used three times throughout the procedure: during Session 1, during Session 2 right before the MAST (but post-intervention and MAST instructions), and after the memory test. The task was described to participants as the (control) MAST and the subsequent memory test. Items in the final time point were written in past tense. The questionnaire consists of 11 questions with responses ranging on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*) and is composed of two subscales: task demands and coping resources ($\alpha_{\text{demand}} = .86$, $\alpha_{\text{resource}} = .78$). The six task demand questions focus on perceptions of danger, uncertainty, and required effort (i.e., "This task is demanding"; "This task is stressful"; "This task is distressing"; "This task is threatening"; "I am uncertain how I will perform"; "This task requires a lot of effort"). The five coping resources questions focus on perceptions of safety and familiarity of the situation, skills, knowledge, and abilities (i.e., "I have the abilities to perform well"; "I have the expectations to perform well"; "Performing well is important to me"; "This task is a positive challenge"; "I am the type of person who does well on these tasks"). Responses were averaged separately for the task demand questions and the coping resources questions, and a threat index was created by calculating the demands/resources ratio.

Mindfulness assessment

Because dispositional mindfulness has been associated with increased coping ability and decreased cortisol reactivity (e.g., Bergomi et al., 2013; Daubenmier et al., 2014), we used the Kentucky Inventory of Mindfulness Skills (KIMS; Baer

et al., 2004) to assess mindfulness techniques. This measure is comprised of 39 items that are rated on a scale from 1 (*Never or very rarely true*) to 5 (*Very often or always true*). It assesses four mindfulness skills: Observing ($\alpha = .82$), Describing ($\alpha = .92$), Acting with awareness ($\alpha = .79$), and Accepting without judgment ($\alpha = .89$). Participants completed this measure during Session 1 before the encoding began.

Procedure

During Session 1, after filling out the informed consent form, participants completed the stress appraisal questionnaire and the KIMS. After a blood pressure measure, they completed the PANAS. Next, participants viewed the images. To ensure engagement and motivation, they were first informed that the top three scorers on the later test would each receive an additional 15-euro voucher upon completion of the study. Session 1 ended with another PANAS. The next day, participants returned to the lab (24–26 h later) for Session 2 and completed a demographic questionnaire as well as other baseline data measures, including subjective stress. After 10 min had passed since they started, participants gave a baseline blood pressure and completed the PANAS. Participants then received the placebo or actual intervention manipulations. Participants were asked to rephrase each summary in writing in their own words to ensure engagement with the task. These tasks took around 10 min to complete. Next, participants completed the PANAS and engaged in the instruction section of the MAST before completing the stress appraisal questionnaire. They next engaged in the (control) MAST. Similar intervals (~5 min) between the intervention and relevant tasks have been used in past research (e.g., Jamieson et al., 2016). Next, participants again completed the subjective stress

questionnaires. After 10 min had passed since the end of the MAST (to assess memory retrieval during the cortisol peak; Smeets et al., 2012), participants began the memory test. First, the experimenter explicitly reminded participants to continue to utilise the information they learned in the summaries of the scientific articles. Partway through data collection, experimenters also began to remind participants to use their native language, because some participants filled out the memory test in the wrong language. Then, participants were tested on their memory for the images using free recall (for two images) and multiple-choice recognition questions (for two images). Finally, participants once again completed subjective questionnaires focusing on subjective stress ratings and the stress appraisal questionnaire. Blood pressure was measured ten times throughout the session. Figure 1 shows an overview of the timeline.

Results

Manipulation checks

To assess the successfulness of the stress induction, we compared subjective stress (PANAS-Positive Affect and Negative Affect scores; six within-levels) and blood pressure (systolic and diastolic; ten within-levels) between the Stress-intervention, Stress-placebo, No-stress-placebo control conditions using a mixed ANOVA. We also computed a mixed ANOVA to examine stress appraisal responses (scored separately for task demands and coping resources and combined into a threat index by dividing the two; three within-levels).

Affect

Higher scores on the positive affect scale of the PANAS reflect higher self-reported positive affect, and higher

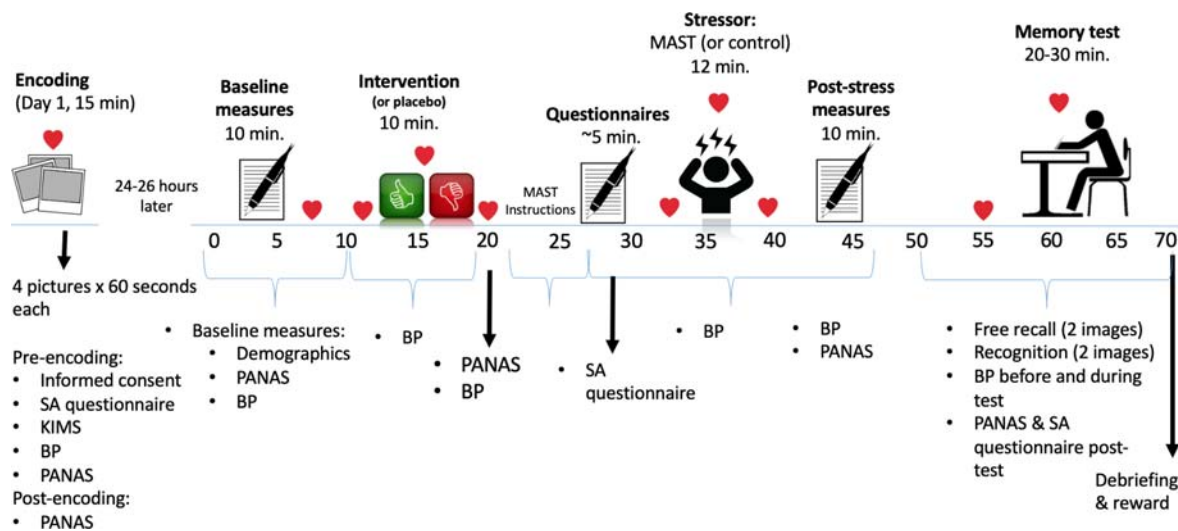


Figure 1. Study timeline and procedure.

Note: SA = stress appraisal. KIMS = Kentucky Inventory of Mindfulness Skills. PANAS = Positive and Negative Affect Schedule. ♥ and BP = blood pressure. MAST = Maastricht Acute Stress Test.

scores on the negative affect scale of the PANAS reflect higher self-reported negative affect. Stress differentially affected negative affect scores depending on timing, $F(6.886, 174) = 14.561, p < .001, \eta_p^2 = .143$. Follow-up tests revealed differences between conditions pre-stressor: during Session 1 pre-encoding, $F(2, 174) = 4.555, p = .012, \eta_p^2 = .050$, Session 1 post-encoding, $F(2, 174) = 3.159, p = .045, \eta_p^2 = .035$, and Session 2 baseline, $F(2, 174) = 5.316, p = .006, \eta_p^2 = .058$. Pairwise comparisons indicated that negative affect scores in the Stress-intervention condition were statistically significantly higher than in the No-stress-placebo control condition during Session 1 pre-encoding ($p = .011$). Likewise, the Stress-intervention condition showed higher negative affect scores than both the No-stress-placebo control and Stress-placebo conditions at Session 2 baseline ($p = .015$ and $p = .017$, respectively). Due to the unexpected initial difference in negative affect between groups in Session 1, we also conducted an exploratory ANCOVA analysis, examining the Session 1 pre-encoding negative affect scores as a covariate. These results showed a significant interaction between stress and timing, $F(6.183, 534.835) = 17.126, p < .001, \eta_p^2 = .165$, though there was no longer a statistically significant difference between conditions in the Session 1 post-encoding and Session 2 baseline negative affect scores.

In addition and as expected, we observed differences post-MAST, $F(2, 174) = 31.292, p < .001, \eta_p^2 = .265$. All conditions differed post-MAST (Stress-intervention vs. Stress-placebo: $p < .001$; Stress-intervention vs. No-stress-placebo control: $p < .001$; Stress-placebo vs. No-stress-placebo control: $p = .001$), with the Stress-intervention condition showing the highest negative affect scores

($M = 20.627, SE = .754$), followed by the Stress-placebo condition ($M = 16.102, SE = .754$), and finally, the No-stress-placebo control condition ($M = 12.203, SE = .754$). For positive affect, we found no statistically significant difference between conditions, $F(8.424, 174) = 1.017, p = .422, \eta_p^2 = .012$. Figure 2 displays changes in negative and positive affect scores over time.

Blood pressure

Stress differentially affected blood pressure depending on the timing for both systolic and diastolic measures, $F(11.622, 174) = 13.157, p < .001, \eta_p^2 = .131$ and $F(12.136, 174) = 15.094, p < .001, \eta_p^2 = .148$, respectively. Follow-up tests revealed no statistically significant differences between conditions at baseline, systolic: $F(2, 174) = 1.238, p = .292, \eta_p^2 = 0.014$, diastolic: $F(2, 174) = .052, p = .949, \eta_p^2 < .001$. However, during the stressor (i.e., MAST), the stress conditions showed statistically significantly higher blood pressure than the non-stressed condition, systolic: $F(2, 174) = 25.208, p < .001, \eta_p^2 = .225$ (Stress-intervention vs. Stress-placebo: $p > .999$; Stress-intervention vs. No-stress-placebo control: $p < .001$; Stress-placebo vs. No-stress-placebo control: $p < .001$), diastolic: $F(2, 174) = 32.672, p < .001, \eta_p^2 = .273$ (Stress-intervention vs. Stress-placebo: $p > .999$; Stress-intervention vs. No-stress-placebo control: $p < .001$; Stress-placebo vs. No-stress-placebo control: $p < .001$). Figure 3 shows changes in blood pressure measurements across time.

Stress appraisal

A threat index was created by calculating the demands/resources ratio from the Stress Appraisal subscales, with numbers >1 representing task demands outweighing

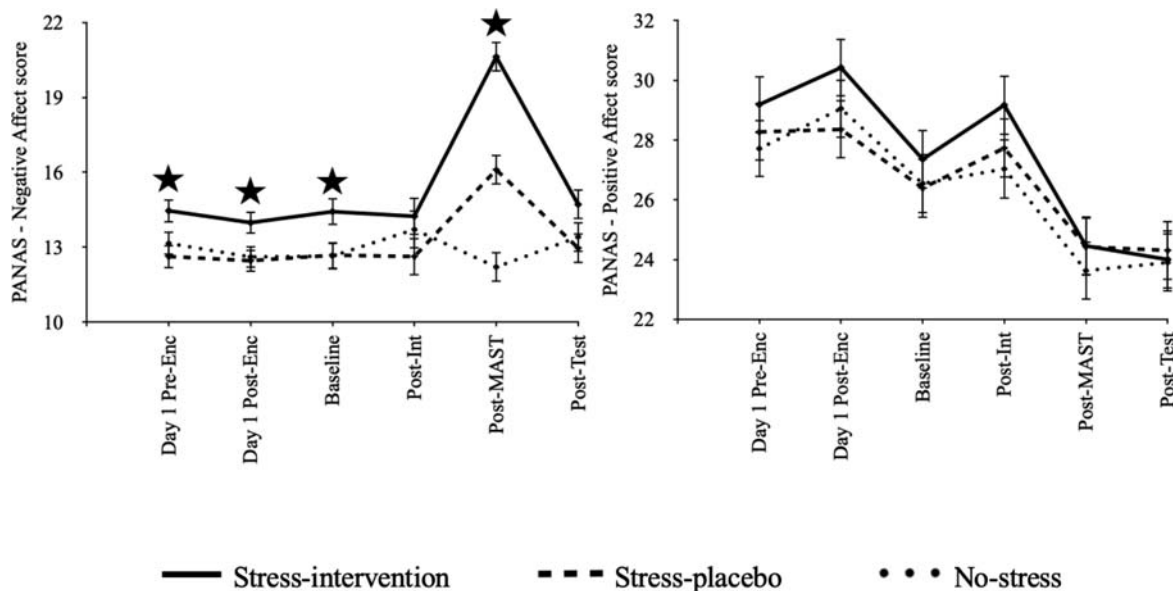


Figure 2. Negative and positive affect scores over time per condition.

Note: Different Y-axis labels for display purposes. PANAS = Positive and Negative Affect Schedule (Watson et al., 1988). Enc = Encoding. Int = Intervention/Placebo. MAST = Maastricht Acute Stress Test. Test = memory test. Error bars = standard error. ★ = $p < .05$. Possible negative and positive affect scores range between 10 and 50.

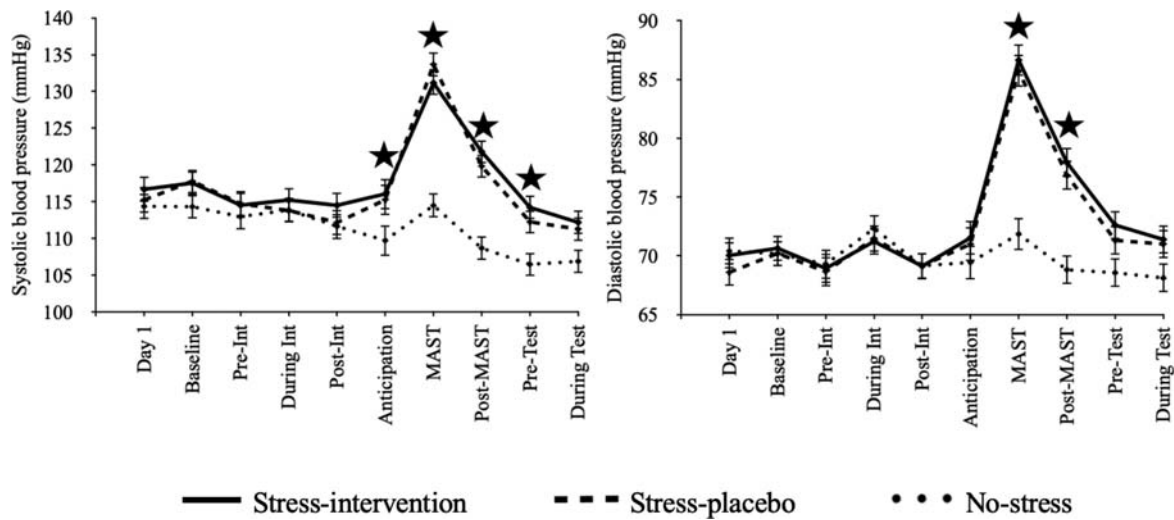


Figure 3. Systolic and diastolic blood pressure over time per condition.

Note: mmHg = millimeters of mercury. Int = Intervention/Placebo. Anticipation = Anticipation phase of the MAST. MAST = Maastricht Acute Stress Test. Test = memory test. Error bars = standard error. ★ = $p < .05$. For each significant indicator, both stress conditions showed higher blood pressures than the no-stress condition, apart from the Anticipation phase for systolic, in which the Stress-intervention condition showed a statistically significant higher systolic blood pressure than the No-stress-placebo control condition ($p = .030$) but the Stress-placebo condition did not ($p = .073$).

coping resources. Stress condition differentially affected the threat index depending on the timing, $F(3.712, 174) = 19.624$, $p < .001$, $\eta_p^2 = .184$. Follow-up tests revealed no statistically significant differences during Session 1, $F(2, 174) = .677$, $p = .510$, $\eta_p^2 = .008$. Pre-MAST, both Stress conditions showed a higher threat index than the No-stress-placebo control condition, $F(2, 174) = 52.255$, $p < .001$, $\eta_p^2 = .375$ (Stress-intervention vs. Stress-placebo: $p > .999$; Stress-intervention vs. No-stress-placebo control: $p < .001$; Stress-placebo vs. No-stress-placebo control: $p < .001$). Additionally, but to a lesser extent, the Stress-intervention condition showed a higher

threat index than the No-stress-placebo control condition Post-test ($p < .001$), $F(2, 174) = 9.775$, $p < .001$, $\eta_p^2 = .101$. Figure 4 shows changes in threat index over time. Table B in supplementary materials shows exploratory analyses of stress effects on task demands and coping resources subscales separately.

Main analyses: effect of stress and intervention on memory performance

We examined the effect of stress and intervention condition on recall quantity (number of total details, correct details,

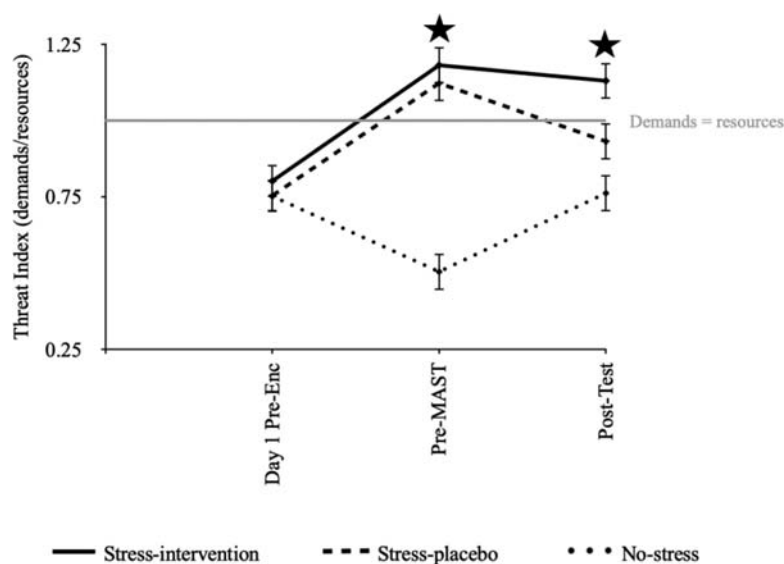


Figure 4. Threat index (demands/resources) across conditions over time.

Note: An index > 1 represents task demands outweighing coping resources. Enc = encoding. MAST = Maastricht Acute Stress Test. Test = memory test. Error bars = standard error. ★ = $p < .05$. At pre-MAST, both Stress conditions statistically significantly differed from the No-stress-placebo control condition ($p < .001$). Post-test, only the Stress-intervention and No-stress-placebo control conditions statistically significantly differed from each other ($p < .001$).

Table 1. Descriptive and inferential statistics for memory performance measures.

	Stress-intervention		Stress-placebo		No-stress-placebo		Stress effects		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	η_p^2
Free recall									
Quantity	70.745	19.340	72.596	23.015	69.365	18.028	.334	.716	.004
Accuracy	.877	.072	.863	.085	.880	.079	.657	.520	.008
Recognition									
Accuracy	.548	.100	.518	.114	.535	.107	1.172	.312	.013
Sensitivity	-.029	.729	-.163	.720	.061	.672	1.499	.226	.017
Response bias	.235	.320	.216	.289	.100	.245	3.838	.023	.042

Note: Reported results are from one-way ANOVAs with condition (3 groups) as between-subjects factor. Free recall quantity = number of total reported details. Free recall accuracy = correct reported details/total reported details. Recognition accuracy = correct answers/all questions. Recognition sensitivity = d' . Recognition response bias = c . Findings that are statistically significant at the $p < .05$ level are in bold.

and incorrect details), recall accuracy, recognition accuracy, and recognition sensitivity in separate one-way ANOVAs.¹ In planned comparisons using two-tailed tests, we compared performances in the Stress-placebo vs. No-stress-placebo control conditions (i.e., Hypothesis 1) and Stress-placebo vs. Stress-intervention conditions (i.e., Hypothesis 2). Eighteen participants completed the free recall task incorrectly, using English rather than their native Dutch or German language ($n = 3$), describing the wrong image ($n = 13$), or failing to provide a free recall response for one of the images ($n = 2$). We excluded these participants from the free recall analyses, leaving $N = 159$ for these analyses and $N = 177$ for recognition analyses.

Descriptives and inferential statistics regarding stress effects on memory measures are reported in Table 1. In contrast with our predictions, we found no statistically significant differences in recall quantity or recall accuracy as a function of condition (Figures 5 and 6). Recognition accuracy and recognition sensitivity (d') also did not differ statistically significantly as a function of condition. As shown in Table 2, planned multiple comparisons for each memory measure revealed no statistically significant differences between any condition. Thus, in line with the overall test, we did not find support for Hypothesis 1 or Hypothesis 2 in the planned multiple comparisons.

In our planned exploratory analysis of response bias (c) where 0 represents no bias, values < 0 represent a liberal

bias, and values > 0 represent a conservative bias, we found a small but statistically significant effect of stress on response bias, $F(2, 176) = 3.838, p = .023, \eta_p^2 = .042$. Planned exploratory multiple comparisons using Tukey HSD (see Table 2) showed a statistically significant difference between the Stress-intervention and the No-stress-placebo control conditions ($p = .031$). The Stress-intervention condition showed a more conservative response bias than the No-stress-placebo control condition, though responding was generally more conservative than liberal across all conditions (i.e., above 0; less likely to choose a definitive answer, more likely to select the “None of the above” option).

Planned exploratory analyses

Dispositional mindfulness moderation analysis

To examine whether participants' dispositional mindfulness moderated the relationship between acute stress and memory performance, we conducted four moderated regression analyses using the PROCESS v.3.4.1 for SPSS (Model 1; Hayes, 2018). We entered Stress condition into the model as the predictor (X ; 1 = Stress-intervention, 2 = Stress-placebo, 3 = No-stress-placebo control), memory measures as the outcome variable (Y ; i.e., recall quantity, recall accuracy, recognition accuracy, recognition sensitivity), and KIMS score (overall $M = 126.508, SD = 13.751$) as the moderator (W).

Table D in supplementary materials presents a summary of these moderator analyses. The overall models neither predicted a significant amount of variance in recall quantity, $F(3, 155) = 1.112, p = .346, R^2 = .021$, nor recall accuracy, $F(3, 155) = .131, p = .942, R^2 = .003$. Similarly, the overall models neither predicted a significant amount of variance in recognition accuracy, $F(3, 173) = .455, p = .714, R^2 = .008$, nor recognition sensitivity, $F(3, 173) = .159, p = .924, R^2 = .003$. Again, neither interaction term was significant. Thus, dispositional mindfulness did not moderate the relationship between acute stress and memory performance.

Bayes factor

As planned, we used JASP version 0.13.1 (JASP Team, 2020) with the default Cauchy prior (.707) to conduct Bayesian independent samples t-tests between the Stress-intervention and the No-stress-placebo control condition on recall quantity,

Table 2. Planned multiple comparisons for memory performance measures.

	Stress-placebo vs. no-stress-placebo control (Hypothesis 1)	Stress- intervention vs. stress-placebo (Hypothesis 2)	Stress-intervention vs. no-stress- placebo control (Exploratory)
	<i>p</i>	<i>p</i>	<i>p</i>
Free recall			
Quantity	.695	.884	.934
Accuracy	.529	.652	.976
Recognition			
Accuracy	.678	.280	.771
Sensitivity	.200	.559	.769
Response bias	.073	.934	.031

Note: Correction = Tukey HSD. Free recall quantity = number of total reported details. Free recall accuracy = correct reported details/total reported details. Recognition accuracy = correct answers/all questions. Recognition sensitivity = d' . Recognition response bias = c . Findings that are statistically significant at the $p < .05$ level are in bold.

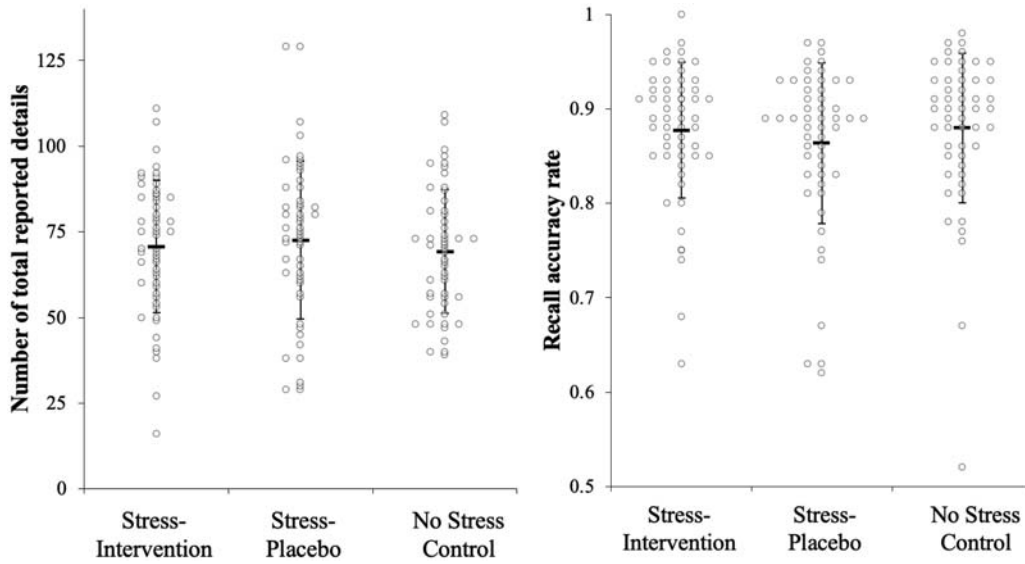


Figure 5. Free recall performance across stress conditions.

Note: =Mean. Errors bars = standard deviation. Figures created using template from Weissgerber et al. (2015).

recall accuracy, recognition accuracy, and recognition sensitivity. We found substantial evidence supporting the null hypothesis across all memory measures (recall quantity: $BF_{01} = 4.578$, error % = 0.018; recall accuracy: $BF_{01} = 4.782$, error % = 0.016; recognition accuracy: $BF_{01} = 4.061$, error % = 0.0004; recognition sensitivity: $BF_{01} = 4.098$, error % = 0.0005; Jarosz & Wiley, 2014; Jeffreys, 1961; Raftery, 1995).

Discussion

In this preregistered experiment, we examined the usefulness of an arousal reappraisal intervention for mitigating negative stress effects on memory performance. We predicted that retrieval stress would impair recognition and recall memory

(Hypothesis 1) and that an arousal reappraisal intervention would mitigate these negative effects of retrieval stress on memory (Hypothesis 2). Although the stressor effectively increased blood pressure and negative affect scores, we did not find statistically significant effects of retrieval stress on recognition or recall memory performance. Furthermore, the arousal reappraisal intervention showed no statistically significant effects on the outcome memory measures. These findings do not support Hypotheses 1 and 2.

Effects of retrieval stress on memory performance

As expected and in line with previous research (e.g., Shilton et al., 2017; Smeets et al., 2012), negative affect and blood

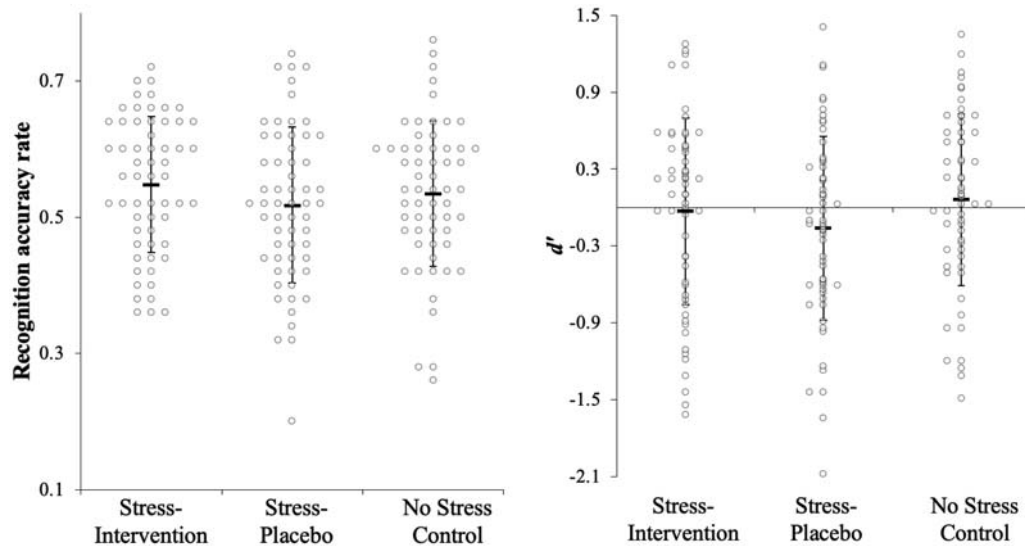


Figure 6. Recognition performance across stress conditions.

Note: =Mean. Errors bars = standard deviation. Figures created using template from Weissgerber et al. (2015).

pressure increased in both stress conditions after experiencing the stressor, but did not increase in the control condition. However, unexpectedly, self-reported negative affect scores were higher in the Stress-intervention group even during Session 1 and at baseline during Session 2. Because participants all experienced the same procedure up until reading the intervention or placebo packets, these results cannot be explained by differences in experimental condition, particularly because our exploratory analyses reveal that the unexpected Session 2 baseline differences were driven by the initial Session 1 pre-encoding differences that took place just a few minutes into Session 1 testing. Regardless of this anomaly, the results confirm the efficacy of the MAST at increasing subjective stress and physiological arousal in the stress conditions but not the control condition.

We found no support for Hypothesis 1, as there were no statistically significant differences between the Stress-placebo and the No-stress-placebo control condition for any of the memory measures. Most published past research demonstrated negative effects of retrieval stress on memory and a recent meta-analysis, examining 26 published papers, also surmised that acute stress at retrieval impaired memory performance (*Hedges' g* = $-.22$; Shields et al., 2017). Still, some experiments only show these impairing effects under certain conditions. For example, past experiments have suggested stronger stress effects for recall as opposed to recognition (de Quervain et al., 2000, 2003; Gagnon & Wagner, 2016). We examined both types of memory test, as research has shown that retrieval stress can affect both types of memory measures (e.g., Kuhlmann et al., 2005; Schönfeld et al., 2014; Schwabe & Wolf, 2014). However, our data show that acute retrieval stress experienced before the memory test did not statistically significantly affect performance for either test type. It is also possible that retrieval stress effects may only be apparent when examining cortisol responders, but not the stress condition as a whole (e.g., Buchanan & Tranel, 2008; Buchanan et al., 2006; Schönfeld et al., 2014). Other research suggests impairing effects in samples of men (e.g., Kuhlmann et al., 2005), but not women (e.g., Hidalgo et al., 2015; Schoofs & Wolf, 2009). Relatedly, meta-analytic effects of retrieval stress on memory performance were only statistically significant when examining studies that excluded women using hormonal contraceptives (Shields et al., 2017).

Some research is more in line with the current findings, showing null effects of retrieval stress on memory. Specifically, stress had no effect on recognition and recall memory of men and women for word lists, autobiographical memories, and videos as stimuli (Becker et al., 2006; Schoofs & Wolf, 2009; Tollenaar et al., 2009; Zoladz et al., 2014). Therefore, research on retrieval stress and memory performance does show mixed findings on related research questions. Thus, although not consistent with our hypotheses, our findings add to existing evidence that effects of acute retrieval stress on memory performance may not be robust under all conditions.

Effects of arousal reappraisal intervention on negative affect and stress appraisals

In addition to examining negative affect as a stress manipulation check, we investigated whether the arousal reappraisal intervention influenced negative affect and stress appraisals. Past research suggested that arousal reappraisal interventions decrease negative affect (e.g., Jamieson et al., 2013a, 2018; Seery, 2011) and increase positive affect (Jentsch & Wolf, 2020). However, our data do not support this idea, with the Stress-intervention condition displaying higher negative affect than even the Stress-placebo condition post-MAST. In fact, we also found that this post-MAST difference between the stress conditions was not driven by the initial group difference at pre-encoding in Session 1, strengthening the idea that the intervention did not reduce negative affect. Our findings do align with past work that similarly showed no change in negative or positive affect for participants engaging in an arousal reappraisal intervention in both a socially evaluated speech task and a painful cold pressor task (Denson et al., 2014). Thus, these findings in sum suggest that arousal reappraisal does not consistently alter positive and negative affect, at least across a variety of laboratory stressors, including the modified TSST, a cold pressor task, and the MAST.

Our findings also do not support the notion that the arousal reappraisal intervention increases coping resources compared to the placebo. In fact, the Stress-intervention condition showed the highest threat indices. Both pre-MAST and post-test, the Stress-intervention condition differed statistically significantly from the control condition and showed a threat index above 1, suggesting that participants felt the task demands outweighed their coping resources. The Stress-placebo condition showed a similar difference with the control condition pre-MAST, but not post-test. Notably, the Stress-intervention condition did not statistically significantly differ from the Stress-placebo condition at any timepoint. These findings are inconsistent with past results suggesting that arousal reappraisal intervention lowers the threat index by increasing self-reported coping resources (Beltzer et al., 2014; Jamieson et al., 2012, 2013b, 2016). However, they are in line with other work showing no such effects of arousal reappraisal on anticipatory stress appraisals (Jentsch & Wolf, 2020). Thus, perhaps the effects of arousal reappraisal interventions on subjective affect and coping resources are not wholly effective across different settings.

Our research differed from previous work examining arousal reappraisal on several few key variables, which may help explain the inconsistent results. The type of stressor used in our experiment may explain the reduced efficacy of the arousal reappraisal intervention (Denson et al., 2009; Liu et al., 2019). The MAST uses a combination of social-evaluative threat and pain-based stress. Specifically, results from a recent meta-analysis showed that arousal reappraisal interventions improved subjective

stress outcomes for active stressors (e.g., those requiring participant engagement, such as performing a speech) but not for passive stressors (e.g., those requiring minimal engagement, such as hand submersion into cold water; Liu et al., 2019). However, other research suggests that reappraising passive, pain-based stressors is possible (Bray, 2015), with participants who engaged in arousal reappraisal displaying increased coping resources when confronted with pain-based stressors (Denson et al., 2014). Still, the passive sections of the MAST may have contributed to the null results observed in the current study. Focusing on changing stress *mindsets* may be better suited for more passive, pain-based tasks, by teaching participants about more general beliefs about the nature of stress rather than manipulating specific appraisals of demands and resources. For example, instructing participants to take on mindsets that stress generally enhances ability can promote adaptive behaviours and less negative outcomes (Crum et al., 2013, 2017; Jamieson et al., 2018). Thus, even when resources are not sufficient to meet the demands of the situation (e.g., the pain of the MAST), improving performance using reappraisal may still be conceivable.

Effects of arousal reappraisal intervention on memory performance

With data suggesting a null effect of acute retrieval stress on memory, we could not examine whether arousal reappraisal can *mitigate* these effects. However, we were still able to test whether the Stress-intervention group showed enhanced memory performance by comparing the stress conditions as planned. We found no statistically significant differences between the Stress-intervention and Stress-placebo condition on any of the memory measures, indicating no support for Hypothesis 2. Additionally, our planned exploratory analyses revealed no evidence that dispositional mindfulness moderated the relationship between acute stress and memory performance, contrary to recent research that showed that greater habitual use of reappraisal leads to more adaptive cardiovascular responses for those engaging in an arousal reappraisal intervention (Jentsch & Wolf, 2020). Finally, we found substantial Bayesian evidence that there was no difference between the Stress-intervention and control condition.

Other previous null findings regarding arousal reappraisal interventions may point to explanations for the current data. Although reappraising arousal can be beneficial for certain cognitive tasks, perhaps it is not useful for all types of cognitive tasks. For example, arousal reappraisal improves executive functioning (Jamieson et al., 2010, 2013a). The authors suggest that perhaps arousal reappraisal benefited quantitative but not qualitative performance in their experiment because math problems rely on more active processing using executive resources than verbal problems. The same idea may apply to various memory

tasks. For example, while free recall requires active recollection, recognition relies on familiarity (Gagnon & Wagner, 2016). As recollection requires greater executive functioning than familiarity (McCabe et al., 2010), perhaps arousal reappraisal would not benefit recognition tasks. However, we neither saw effects of stress nor the intervention on recall memory performance, where executive functioning and active processing play a larger role, suggesting that this hypothesis cannot fully explain our data. Past research has shown enhanced cognitive benefits of arousal reappraisal in math exams (Jamieson et al., 2016), quantitative portions of standardised tests (Jamieson et al., 2010), and public speaking tasks (Beltzer et al., 2014). It is possible that our participants did not have pre-existing negative experiences and appraisals of these familiar contexts (e.g., test anxiety), that participants in other studies had. Of note, arousal reappraisal interventions have not shown cognitive benefits on other types of tasks, including qualitative sections of standardised tests (Jamieson et al., 2010) or math performance competitions against another individual (Hangen et al., 2019).

Limitations and future directions

There were several limitations to this initial investigation into using an arousal reappraisal intervention with a memory performance task. A broader range of physiological measures, including cortisol levels and cardiovascular responses, would provide deeper insight into the results. First, assessing cortisol levels would confirm HPA axis activation, and allow for comparisons between conditions, as arousal reappraisal participants may show more moderate HPA activation than placebo participants (e.g., Jamieson et al., 2018; Jentsch & Wolf, 2020; Mendes & Park, 2014; but see Denson et al., 2014). Additionally, cortisol measures would allow separate examinations of high cortisol responders versus low cortisol responders. This comparison could reveal interesting effects, as some research on retrieval stress and memory suggests that impairing effects are only observed in high cortisol responders (e.g., Buchanan & Tranel, 2008; Buchanan et al., 2006; Schönfeld et al., 2014). In future work, it may also be beneficial to explore the role of individual differences in subjective stress responding and coping (e.g., levels of social anxiety; Jamieson et al., 2013b). Second, although our measures of blood pressure confirmed physiological arousal, blood pressure is an ambiguous measure in terms of determining challenge or threat states (Seery, 2011). Comparing conditions on cardiovascular responses (e.g., heart rate, vascular conductance, cardiac output, and total peripheral resistance) would provide better insight into the challenge versus threat states from a physiological perspective, rather than solely relying on self-reported affect and stress appraisals.

Some recent research also suggests sex effects, such that arousal reappraisal may improve physiological functioning for males but not for females (Hangen et al.,

2019). Sex has also been shown to influence the relationship between retrieval stress and memory (e.g., Cahill, 2012), such that males often show higher cortisol responses to stress than females (e.g., Hidalgo et al., 2019). Hormonal contraceptive use and menstrual cycle also play a role in the stress response (e.g., Kirschbaum et al., 1999; Shields et al., 2017). Female participants using hormonal contraceptives can exhibit blunted stress hormone responses (e.g., Nielsen et al., 2013), leading some experiments to only include females who do not use hormonal contraceptives (e.g., Jentsch & Wolf, 2020). Our sample included mostly females (81.92%), half of whom were using hormonal contraceptives and half of whom were not. Although we partially controlled for variations in sex and hormonal contraceptive usage by balancing these demographic factors across conditions, it is possible that the null effects stem from differences in sex. Data from the current experiment do not allow for a sufficiently powered analysis of potential sex differences, but future studies could treat sex as a factor in the design to allow for a well-powered comparison.

Clearer task instructions for participants may benefit future research in a few ways. First, although participants were instructed to answer the stress reappraisal questionnaire for the task as a whole (i.e., including the MAST and memory test), it is possible that the vague terminology of “the task” in the questionnaire may have confused some participants. This does not appear to have been a problem, as participants show no differences at baseline, but only after they found out to which stress condition they were randomly assigned. Still, future research using this questionnaire could use more precise language to indicate to participants exactly which task they should assess when reporting their levels of task demands and coping resources. Second, free recall responses for eighteen participants were excluded because they completed one of the free recall questions for the wrong scene, did not use their native language, or did not provide a free recall response for one of the images. Memory tasks in future experiments could use a singular to-be-remembered item or a more specific prompt to avoid such confusion.

Some of the data in this experiment also suggest the possibility of a floor effect. Floor effects may make it difficult to observe potential effects of stress and reappraisal on memory performance, as they indicate that participants are performing at very low levels. In the current study specifically, the d' scores, which are around 0, suggest possible floor events for the recognition data. Such floor effects could obscure observation of the predicted negative effects of retrieval stress on sensitivity. These d' values likely stem from the fact that participants were required to make a choice for each recognition multiple choice question thereby increasing false alarm rates. However, the accuracy data do not indicate floor effects, with participants performing well above chance levels. Task difficulty likely played a role in this pattern of

results. Past research has used these materials with very short retention intervals (i.e., 10 s; Sauer & Hope, 2016). Furthermore, although prior experiments investigating retrieval stress effects on memory have used other types of pictures (e.g., Hidalgo et al., 2015), the images used in this study are more complex. We used these stimuli with a 24-hour retention interval to better reflect eyewitness settings, where retention intervals are oftentimes lengthy and experiences complex. Future research might find different results with shorter retention intervals or a less demanding memory task, particularly in regard to measures of sensitivity.

Finally, it is possible that effects of retrieval stress on memory are small rather than medium. We deliberately based our a priori power analysis on substantial previous relevant work suggesting medium effect sizes of retrieval stress effects on memory performance (e.g., Domes et al., 2004; Hidalgo et al., 2015; Kuhlmann et al., 2005; Schwabe & Wolf, 2009, 2014, Experiment 2; Smeets, 2011; Tollenaar et al., 2009). Still, future work should aim to replicate the current findings, perhaps with larger sample sizes to capture even small effects of retrieval stress on memory performance.

There are several interesting future avenues for this line of research. First, examining the efficacy of stress mindsets (i.e., more broadly instructing participants that stress generally enhances ability) on similar stressors could elucidate whether arousal reappraisal is limited to certain types of stressors (i.e., social/evaluative vs. pain, etc.). Second, future experiments could examine memory using other stressful scenarios, where the stress induction is directly relevant to the memory task. For example, in one stress study, participants encoded neutral and emotional pictures and words (Schönfeld et al., 2014). Twenty-four hours later, they encountered either a stressful, oral examination test or a non-stressful free recall test. Using a stressful oral examination aligns more closely with past research on arousal reappraisal (i.e., exam performance, public speaking tasks, etc.), and understanding the efficacy of arousal reappraisal in such a setting would be relevant for certain educational settings (e.g., oral exams, presentations). In addition, such an approach would be a building block for the forensic field, more closely reflecting applied police interviews for witnesses. Third, research could directly compare cognitive tasks already demonstrated to be enhanced by arousal reappraisal (e.g., math questions) with motivated memory tasks to examine potential boundaries of arousal reappraisal effects. That is, a direct comparison would allow for a more informed discussion regarding the scope of the effects of arousal reappraisal on distinctive types of cognitive tasks. Finally, in our exploratory analyses, we found that the Stress-intervention condition showed a more conservative response bias than the No-stress-placebo control condition. Though no past research has directly examined the effects of acute retrieval stress and cognitive appraisals on response bias, this exploratory finding highlights a

route for some interesting future research. For example, future experiments might more closely investigate how arousal reappraisal interventions alter individuals' decision criteria, particularly while experiencing stress prior to or during retrieval.

Conclusion

In conclusion, we found no support for our hypotheses, as acute stress experienced prior to retrieval did not impair recognition and recall memory. Our arousal reappraisal intervention also did not affect memory performance. In light of these findings, we can neither make conclusive statements about the effects of retrieval stress on memory performance beyond our data, nor definitively judge the efficacy of the intervention in all contexts of acute stress and memory performance. Still, given the body of research suggesting negative effects of retrieval stress on memory performance (e.g., Shields et al., 2017; Wolf, 2017), it is important for future research to continue exploring ways to mitigate negative effects of stress, both in terms of cognitive benefits but also in relation to emotional distress and coping ability.

Note

1. In light of the unexpected baseline differences in negative affect between conditions, we also ran ANCOVAs for each main analysis with baseline negative affect as a covariate. Controlling for baseline negative affect did not alter results; these additional analyses are reported in Table C in supplementary materials.

Acknowledgements

We gratefully acknowledge the help of Tessa Bacilio, Emilis Čepulis, Claudia Hekkert, Pauline Lei de Backker, Melissa Nickel, Brandon Smalls, and Michelle Wieberneit during data collection and recruitment of participants.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

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-EPJD to Carey Marr.

Data availability statement

Anonymised data and study materials have been made publicly available on OSF and can be accessed with the following link: https://osf.io/8a9sq/?view_only=ce42932ad0cb415f9934b64b1d7928bf

ORCID

Carey Marr  <http://orcid.org/0000-0002-0665-1811>
 Melanie Sauerland  <http://orcid.org/0000-0002-7401-1365>
 Henry Otgaar  <http://orcid.org/0000-0002-2782-2181>
 Conny W.E.M. Quaedflieg  <http://orcid.org/0000-0003-1769-5438>
 Lorraine Hope  <http://orcid.org/0000-0002-0012-8165>

References

- Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Assessment of mindfulness by self-report: The Kentucky inventory of mindfulness skills. *Assessment, 11*(3), 191–206. <https://doi.org/10.1177/1073191104268029>
- Becker, V. E., Tucker, D. M., Delville, Y., & Mohr, D. C. (2006). Stress facilitates consolidation of verbal memory for a film but does not affect retrieval. *Behavioral Neuroscience, 120*(3), 518–527. <https://doi.org/10.1037/0735-7044.120.3.518>
- Beltzer, M. L., Nock, M. K., Peters, B. J., & Jamieson, J. P. (2014). Rethinking butterflies: The affective, physiological, and performance effects of reappraising arousal during social evaluation. *Emotion, 14*(4), 761–768. <https://doi.org/10.1037/a0036326>
- Bergomi, C., Ströhle, G., Michalak, J., Funke, F., & Berking, M. (2013). Facing the dreaded: Does mindfulness facilitate coping with distressing experiences? A moderator analysis. *Cognitive Behaviour Therapy, 42*(1), 21–30. <https://doi.org/10.1080/16506073.2012.713391>
- Blascovich, J. (2008). Challenge and threat. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation* (pp. 431–445). Psychology Press.
- Blascovich, J., & Mendes, W. B. (2010). Social psychophysiology and embodiment. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *The handbook of social psychology* (5th ed., pp. 194–227). Wiley.
- Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. In M. P. Zanna (Ed.), *Advances in experimental social psychology, Vol. 28* (pp. 1–51). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60235-X](https://doi.org/10.1016/S0065-2601(08)60235-X)
- Bray, N. (2015). Reappraising pain. *Nature Reviews Neuroscience, 16*(3), 124–125. <https://doi.org/10.1038/nrn3919>
- 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>
- 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>
- 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>
- Crum, A. J., Akinola, M., Martin, A., & Fath, S. (2017). The role of stress mindset in shaping cognitive, emotional, and physiological responses to challenging and threatening stress. *Anxiety, Stress, & Coping, 30*(4), 1–17. <https://doi.org/10.1080/10615806.2016.1275585>
- Crum, A. J., Salovey, P., & Achor, S. (2013). Rethinking stress: The role of mindsets in determining the stress response. *Journal of Personality and Social Psychology, 104*(4), 716–733. <https://doi.org/10.1037/a0031201>
- Daubenmier, J., Hayden, D., Chang, V., & Epel, E. (2014). It's not what you think, it's how you relate to it: Dispositional mindfulness moderates the relationship between psychological distress and the

- cortisol awakening response. *Psychoneuroendocrinology*, 48, 11–18. <https://doi.org/10.1016/j.psychoneu.2014.05.012>
- Denson, T. F., Creswell, J. D., Terides, M., & Blundell, K. (2014). Cognitive reappraisal increases neuroendocrine reactivity to acute social stress and physical pain. *Psychoneuroendocrinology*, 49(1), 69–78. <https://doi.org/10.1016/j.psychoneu.2014.07.003>
- Denson, T. F., Spanovic, M., & Miller, N. (2009). Cognitive appraisals and emotions predict cortisol and immune responses: A meta-analysis of acute laboratory social stressors and emotion inductions. *Psychological Bulletin*, 135(6), 823–853. <https://doi.org/10.1037/a0016909>
- 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 blood 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>
- 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>
- Hangen, E. J., Elliot, A. J., & Jamieson, J. P. (2019). Stress reappraisal during a mathematics competition: Testing effects on cardiovascular approach-oriented states and exploring the moderating role of gender. *Anxiety, Stress, and Coping*, 32(1), 95–108. <https://doi.org/10.1080/10615806.2018.1530049>
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis* (2nd ed.). The Guilford Press.
- Hidalgo, V., Pulpulos, M. M., Puig-Perez, S., Espin, L., Gomez-Amor, J., & Salvador, A. (2015). Acute stress affects free recall and recognition of pictures differently depending on age and sex. *Behavioral Brain Research*, 292, 393–402. <https://doi.org/10.1016/j.bbr.2015.07.011>
- Hidalgo, V., Pulpulos, 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>
- Jamieson, J. P., Crum, A. J., Goyer, J. P., Marotta, M. E., & Akinola, M. (2018). Optimizing stress responses with reappraisal and mindset interventions: An integrated model. *Anxiety, Stress, & Coping*, 31(3), 245–261. <https://doi.org/10.1080/10615806.2018.1442615>
- Jamieson, J. P., Mendes, W. B., Blackstock, E., & Schmader, T. (2010). Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE. *Journal of Experimental Social Psychology*, 46(1), 208–212. <https://doi.org/10.1016/j.jesp.2009.08.015>
- Jamieson, J. P., Mendes, W. B., & Nock, M. K. (2013a). Improving acute stress responses: The power of reappraisal. *Current Directions in Psychological Science*, 22(1), 51–56. <https://doi.org/10.1177/0963721412461500>
- Jamieson, J. P., Nock, M. K., & Mendes, W. B. (2012). Mind over matter: Reappraising arousal improves cardiovascular and cognitive responses to stress. *Journal of Experimental Psychology: General*, 141(3), 417–422. <https://doi.org/10.1037/a0025719>
- Jamieson, J. P., Nock, M. K., & Mendes, W. B. (2013b). Changing the conceptualization of stress in social anxiety disorder: Affective and physiological consequences. *Clinical Psychological Science*, 1(4), 1–12. <https://doi.org/10.1177/2167702613482119>
- Jamieson, J. P., Peters, B. J., Greenwood, E. J., & Altose, A. J. (2016). Reappraising stress arousal improves performance and reduces evaluation anxiety in classroom exam settings. *Social Psychology and Personality Science*, 7(6), 579–587. <https://doi.org/10.1177/1948550616644656>
- 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.
- Jentsch, V. L., & Wolf, O. T. (2020). The impact of emotion regulation on cardiovascular, neuroendocrine and psychological stress responses. *Biological Psychology*, 154, 107893. <https://doi.org/10.1016/j.biopsycho.2020.107893>
- 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>
- 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>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- 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>
- Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, 46(8), 819–834. <https://doi.org/10.1037/0003-066X.46.8.819>
- Leitenberg, H. (1990). *Handbook of social and evaluation anxiety*. Plenum Press.
- 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>
- Liu, J. J. W., Ein, N., Gervasio, J., & Vickers, K. (2019). The efficacy of stress reappraisal interventions on stress responsivity: A meta-analysis and systematic review of existing evidence. *PLoS ONE*, 14(2), e0212854. <https://doi.org/10.1371/journal.pone.0212854>
- McCabe, D. P., Roediger, H. L. III, McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, 24(2), 222–243. <https://doi.org/10.1037/a0017619>
- McCarthy, J., & Goffin, R. (2004). Measuring job interview anxiety: Beyond weak knees and sweaty palms. *Personnel Psychology*, 57(3), 607–637. <https://doi.org/10.1111/j.1744-6570.2004.00002.x>
- Mendes, W. B., Gray, H. M., Mendoza-Denton, R., Major, B., & Epel, E. S. (2007). Why egalitarianism might be good for your health: Physiological thriving during stressful intergroup encounters. *Psychological Science*, 18(11), 991–998. <https://doi.org/10.1111/j.1467-9280.2007.02014.x>
- Mendes, W. B., & Park, J. (2014). Neurobiological concomitants of motivational states. In A. J. Elliot (Ed.), *Advances in motivation*

- science (pp. 233–270). Elsevier. <https://doi.org/10.1016/bs.adms.2014.09.001>
- 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. <https://doi.org/10.1016/j.biopsycho.2012.10.007>
- Posthuma, R. A., Morgeson, F. P., & Campion, M. A. (2002). Beyond employment interview validity: A comprehensive narrative review of recent research and trends over time. *Personnel Psychology*, 55(1), 1–81. <https://doi.org/10.1111/j.1744-6570.2002.tb00103.x>
- Quaedflieg, C. W. E. M., & Schwabe, L. (2018). Memory dynamics under stress. *Memory (Hove, England)*, 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>
- Sauer, J., & Hope, L. (2016). The effects of divided attention at study and reporting procedure on regulation and monitoring for episodic recall. *Acta Psychologica*, 169, 143–156. <https://doi.org/10.1016/j.actpsy.2016.05.015>
- 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.psyneuen.2013.09.020>
- Schoofs, D., & Wolf, O. T. (2009). Stress and memory retrieval in women: No strong impairing effect during the luteal phase. *Behavioral Neuroscience*, 123(3), 547–554. <https://doi.org/10.1037/a0015625>
- Schwabe, L., & Wolf, O. T. (2009). Stress prompts habit behavior in humans. *Journal of Neuroscience*, 29(22), 7191–7198. <https://doi.org/10.1523/JNEUROSCI.0979-09.2009>
- 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>
- Seery, M. D. (2011). Resilience: A silver lining to experiencing adverse life events? *Current Directions in Psychological Science*, 20(6), 390–394. <https://doi.org/10.1177/0963721411424740>
- Shields, G. S. (2020). Stress and cognition: A user's guide to designing and interpreting studies. *Psychoneuroendocrinology*, 112, 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>
- Shilton, A. L., Laycock, R., & Crewther, S. G. (2017). The Maastricht Acute Stress Test (MAST): Physiological and subjective responses in anticipation, and post-stress. *Frontiers in Psychology*, 8(567), 1–10. <https://doi.org/10.3389/fpsyg.2017.00567>
- Smeets, T. (2011). Acute stress impairs memory retrieval independent of time of day. *Psychoneuroendocrinology*, 36(4), 495–501. <https://doi.org/10.1016/j.psyneuen.2010.08.001>
- 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., Floerke, V. A., & Thomas, A. K. (2016). Retrieval practice protects memory against acute stress. *Science*, 354(6315), 1046–1048. <https://doi.org/10.1126/science.aah5067>
- 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>
- Sydesman, S. J., Cascardi, M., Poythress, N. G., & Ritterbrand, L. M. (1997). Procedural justice in the context of civil commitment: A critique of Tyler's analysis. *Psychology, Public Policy & Law*, 3(1), 207–221. <https://doi.org/10.1037/1076-8971.3.1.207>
- Tollenaar, M. S., Elzinga, B. M., Spinhoven, P., & Everaerd, W. (2009). Autobiographical memory after acute stress in healthy young men. *Memory (Hove, England)*, 17(3), 301–310. <https://doi.org/10.1080/09658210802665845>
- 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), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Weisserger, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond bar and line graphs: Time for a new data presentation paradigm. *PLoS Biology*, 13(4), e1002128. <https://doi.org/10.1371/journal.pbio.1002128>
- Wolf, O. T. (2017). Stress and memory retrieval: Mechanisms and consequences. *Current Opinions in Behavioral Sciences*, 14, 40–46. <https://doi.org/10.1016/j.cobeha.2016.12.001>
- 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.psyneuen.2014.06.01>