Research Article

Expectancy Violation Drives Memory Boost for Stressful Events

Felix Kalbe, Stina Bange, Annika Lutz, and Lars Schwabe
Department of Cognitive Psychology, Institute of Psychology, Universität Hamburg

Abstract

Stressful events are often vividly remembered. Although generally adaptive to survival, this emotional-memory enhancement may contribute to stress-related disorders. We tested here whether the enhanced memory for stressful events is due to the expectancy violation evoked by these events. Ninety-four men and women underwent a stressful or control episode. Critically, to manipulate the degree of expectancy violation, we gave participants either detailed or minimal information about the stressor. Although the subjective and hormonal stress responses were comparable in informed and uninformed participants, prior information about the stressor abolished the memory advantage for core features of the stressful event, tested 7 days later. Using functional near-infrared spectroscopy, we further linked the expectancy violation and memory formation under stress to the inferior temporal cortex. These data are the first to show that detailed information about an upcoming stressor and, by implication, a reduced expectancy violation attenuates the memory for stressful events.

Emotionally arousing events are typically much better remembered than mundane events. For instance, we are more likely to remember a surprise birthday party or a heated discussion than what we had for lunch last Wednesday. Extensive research confirms that information encoded under stress or states of heightened emotional arousal is preferentially stored in memory (Joëls, Pu, Wiegert, Oitzl, & Krugers, 2006; Sandi, Loscertales, & Guaza, 1997; Vogel & Schwabe, 2016; Wiemers, Sauvage, Schoofs, Hamacher-Dang, & Wolf, 2013). It has further been shown that there is a memory advantage in particular for the central features of emotional or stressful events (Wiemers et al., 2013), whereas the memory for peripheral, less salient elements of these events may even be reduced (Kensinger, Garoff-Eaton, & Schacter, 2007). The superior memory for emotional events is generally adaptive because it helps us to prepare for similar situations in the future. However, overly strong emotional memory may be maladaptive and contribute to anxiety disorders or posttraumatic stress disorder (PTSD; de Quervain, Schwabe, & Roozendaal, 2017; Pitman et al., 2012).

Traditionally, the memory boost for emotional events has been attributed to the many hormones and neurotransmitters that are released during these arousing events, such as glucocorticoids (cortisol in humans) and noradrenaline (Joëls et al., 2006; Schwabe, Joëls, Roozendaal, Wolf, & Oitzl, 2012). These substances may act directly on brain areas critical for memory formation (de Quervain et al., 2017; Joëls et al., 2006). In addition, noradrenaline has been shown to promote a shift toward a salience network, which includes areas such as the amygdala, the temporoparietal junction (TPJ), and the inferior temporal cortex and leads to the prioritized processing and storage of the most salient elements of an ongoing event (Hermans, Henckens, Joëls, & Fernandez, 2014). Although the physiological, arousal-based model of emotional-memory formation is well established, it focuses less on the cognitive mechanisms involved in

Corresponding Author:
Lars Schwabe, Universität Hamburg, Institute of Psychology, Department of Cognitive Psychology
E-mail: lars.schwabe@uni-hamburg.de

memory formation for emotional or stressful events. Recently, we proposed an alternative model that assumes that emotionally arousing events are characterized by their unpredictability and are therefore linked to a prediction error or expectancy violation (Trapp, O'Doherty, & Schwabe, 2018). For instance, a surprise birthday party or heated discussion is less predictable and deviates more strongly from our experience than a common lunch at work. A prediction error or expectancy violation is thought to be a driving force for new learning and plasticity, both at encoding (Greve, Cooper, Kaula, Anderson, & Henson, 2017; Rescorla & Wagner, 1972) and after memory reactivation, during a window of reconsolidation (Sevenster, Beckers, & Kindt, 2013; Sinclair & Barense, 2019). Accordingly, the superior memory for arousing events may be due to the prediction error or expectancy violation associated with these rather unusual events. This model leads to the intriguing hypothesis that the memory boost for emotional events can be attenuated or even prevented by reducing the expectancy violation evoked by these events. Such a mechanism would have important implications for contexts in which individuals may suffer from painful memories, whether on the battlefield or in the emergency room.

Thus, we aimed here to test, for the first time, whether the enhanced memory for a stressful event can be reduced by modulating the expectancy violation associated with this event. To this end, we exposed healthy participants to a standardized stress or control manipulation, including several contextual details to create a rich episode. Critically, whereas some participants received only minimal information before they underwent the rather unusual stressor (or control procedure), which was therefore associated with a strong expectancy violation, others received detailed information about the upcoming event, which we expected would minimize the associated expectancy violation. During the stressful (or control) encounter, we measured brain activity from cortical areas of the salience network, from cortical areas of an executive-control network, and from visual and motor control areas using functional near-infrared spectroscopy (fNIRS). In contrast to functional MRI (fMRI), fNIRS has the great advantage that cortical activity can be recorded in less artificial, real-life situations. Memory for the stressful (or control) episode was tested after 1 week. We hypothesized that detailed prior information would reduce the superior memory for central features of the stressful event. On the brain level, we focused on the inferior temporal cortex, which integrates attention, visual processing, and memory (Miyashita, 1993); is part of the salience network (Hermans et al., 2014); and has been linked to memory formation under stress in a previous fMRI

Statement of Relevance

Stress is ubiquitous in our everyday lives and has a major impact on our health and well-being. Moreover, stress may change the way we think, learn, and remember. For instance, it is well known that stressful events are much better remembered than mundane events, which may have important implications for memory of traumatic events. In this research, we investigated whether the superior memory for stressful events could be alleviated by informing individuals about the upcoming event in as much detail as possible beforehand. We found that prior information indeed attenuated the enhanced memory for the stressful encounter. Moreover, prior information about the upcoming stressor reduced the brain response to the stressful event. These findings provide novel insights into the mechanisms involved in memory formation under stress and may be helpful for developing new approaches to prevent the emergence of debilitating memories, for instance, in military or emergency contexts.

study (Henckens, Hermans, Pu, Joëls, & Fernandez, 2009). We predicted that inferior temporal activity under stress would be attenuated if the stressor-related expectancy violation was low.

Method

Participants

Ninety-seven healthy volunteers (51 female) participated in the experiment (age: $M = 24.84$ years, $SD = 4.05$, range = 18–36 years). This sample size was based on those used in previous studies on stress and memory in our lab (Schwabe, Bohringer, Chatterjee, & Schachinger, 2008; Vogel & Schwabe, 2016) and an a priori power calculation using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007). The power analysis indicated that a sample of 96 would be sufficient to detect a medium-size effect ($f = .20$) in a mixed-design analysis of variance (ANOVA) with a power of .90. All participants were fluent German speakers, had no history of any psychiatric or neurological disorder, had no acute illness, did not take any prescribed medication, and had no background in psychology. Moreover, smokers and women taking hormonal contraceptives were excluded from participation because these factors may affect the endocrine stress response. Three participants were excluded from the analyses: one because he suffered from severe headache due to the fNIRS measurement and
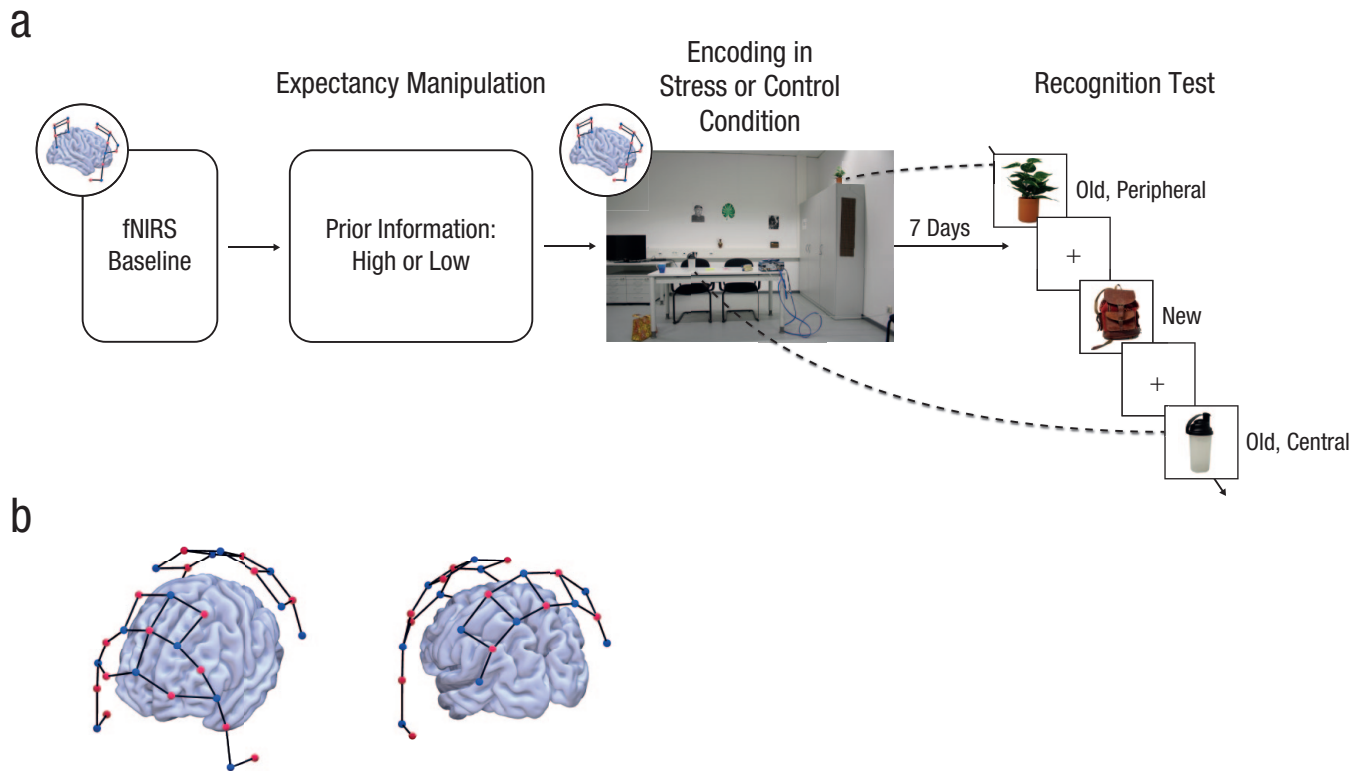
a

b

Fig. 1. Overview of the experimental procedure. Participants underwent a psychosocial stressor or a nonstressful control manipulation with a two-experimenter panel (a). Critically, immediately before the stressful or control encounter, we gave participants either minimal or detailed information about the upcoming treatment, thus manipulating the degree of expectancy violation associated with the treatment. During both manipulations, the activity of predefined cortical areas was measured using functional near-infrared spectroscopy (fNIRS). One week later, participants completed an old/new recognition test for central and peripheral items that had appeared in the stressful or control encounter. Central items were cues with which the panel members interacted and that were therefore part of the stress manipulation, whereas cues that were not part of the stress manipulation were defined as peripheral items. The fNIRS montage (b; sources are shown in red, detectors in blue) covered areas that belong to the salience network recruited under stress, areas implicated in higher order cognition, and sensory and motor control areas.

two because they were identified as outliers in memory performance (3 *SD* below the group mean; Tabachnik & Fidell, 2001). Thus, the final sample consisted of 94 participants. All participants provided written informed consent before participation and received monetary compensation. The study was carried out in accordance with the provisions of the World Medical Association Declaration of Helsinki and approved by the local ethics committee.

Materials and procedure

Participants were tested individually on 2 days, with an interval of 1 week (Fig. 1). All testing took place in the afternoon to control for the diurnal rhythm of the stress hormone cortisol. After providing written informed consent on Day 1 of the experiment, participants completed the Beck Depression Inventory (Beck & Steer, 1987), the State-Trait Anxiety Inventory (Spielberger,

Gorsuch, & Luchene, 1970), and the Trier Inventory for Chronic Stress (Schulz, Schlotz, & Becker, 2004) to control for group differences in depressive mood, anxiety, and chronic stress level. Next, we collected baseline measurements of subjective mood as well as a first saliva sample, and participants underwent a 5-min baseline fNIRS session, during which they were standing in a quiet room. In a fully crossed, between-subjects design with the factors treatment (stress vs. control) and expectancy violation (prior information vs. no prior information), each participant was then randomly assigned to one of four experimental groups: informed stress (*n* = 24), uninformed stress (*n* = 24), informed control (*n* = 23), and uninformed control (*n* = 23).

Day 1: encoding under stress. In the stress condition, participants were exposed to a modified version of the Trier Social Stress Test (TSST; Kirschbaum, Pirke, & Hellhammer, 1993), a gold standard in experimental stress

research that leads to robust subjective and physiological stress responses. In brief, the TSST consisted of a mock job interview during which participants were asked to give a 5-min free speech about why they are the ideal candidate for a job tailored to their interests and to perform a 5-min mental arithmetic task (counting backward in steps of 17 from 2,043). Both the free-speech task and the mental-arithmetic task were performed in front of a rather cold and nonreinforcing panel of two experimenters who were dressed in white coats and introduced as experts in "behavioral analysis." In addition, participants were videotaped throughout the TSST and could see themselves on a large screen placed behind the panel. We modified the standard TSST procedure by incorporating a predefined sequence of actions during which the panel members interacted with specific items (e.g., checking a cell phone, sharpening a pencil). In addition, we provided a number of contextual cues in the experimental room (e.g., a bag, a plant; Fig. 1). Cues with which the panel members interacted and that were therefore part of the stress manipulation were defined as *central items*, whereas cues that were not part of the stress manipulation were defined as *peripheral items.*

In the control condition, participants interacted with two experimenters in a nonstressful way. They first had a 5-min conversation with the experimenters about a topic of the participants' choice (e.g., a book they had just read or a movie they had seen), followed by a simple 5-min counting game. During the control manipulation, the experimenters behaved in a normal, friendly manner and were not dressed in white lab coats, and participants were not videotaped. Importantly, the experimenters performed the same set of actions as during the stress manipulation at the exact same time points relative to the beginning of the manipulation, and the control procedure took place in the same room as the TSST. Thus, the central and peripheral items were identical in the two protocols, with the critical difference that encoding took place under stressful or nonstressful conditions. Notably, the TSST or control procedure took place in a different room from all other assessments before and after the treatment; thus, the encoding in this context was limited to the time of the TSST and control manipulations. Brain activity was recorded throughout the stress or control procedure using fNIRS (see below).

To assess the effective stress induction through the (modified) TSST, we took subjective and endocrine measures at several time points before and after the experimental manipulation. Participants completed a German mood questionnaire (Steyer, Schwenkmezger, Notz, & Eid, 1994) before and after the TSST and control manipulations. Furthermore, they rated the stressfulness, unpleasantness, and difficulty of the treatment on a scale from 1 (*not at all*) to 10 (*very much*) immediately after the end of the TSST and control manipulations. Finally, we analyzed the concentration of the stress hormone cortisol with a luminescence assay (IBL International, Hamburg, Germany) from saliva samples that were collected before and immediately after the TSST and control manipulations, as well as 30 and 45 min after the onset of the experimental manipulation. Because of experimenter error, the saliva samples of 37 participants were lost. However, the number of participants with available cortisol data was comparable across groups (informed stress: $n = 14$, uninformed stress: $n = 13$, informed control: $n = 15$, uninformed control: $n = 15$).

Day 1: manipulation of expectancy. To test the role of expectancy violation in memory formation under stress, we manipulated participants' expectancy of the treatment immediately before they entered the stress and control manipulations. More specifically, whereas half of the stress group received the standard TSST instruction, which mainly involves the notion that participants are requested to give a free speech and that they will receive further information from the panel, the other half of the stress group was informed about the test procedure in as much detail as possible immediately before entering the room where the TSST took place. The informed group was told about the video recordings; the screen on which they could see themselves; the cold, nonreinforcing behavior of the panel; the fact that the panel was dressed in white lab coats and constantly taking notes; and all tasks and procedures. Likewise, half of the control group was informed in detail about the control procedure, whereas the other half was informed only about the conversation at the beginning and that they would receive further information from the other experimenters. Thus, we experimentally manipulated the information that participants received before the beginning of the treatment, which would, however, most likely result in a modification of the expectancy violation that is elicited by the psychosocial stressor, which deviates from typical social interactions in everyday life.

To assess whether the prior information did indeed reduce the expectancy violation evoked by the rather unusual TSST (or control) situation, we had participants complete a questionnaire after the treatment, in which they indicated to what extent the previous situation was surprising, unexpected, predicted, and in line with their prior expectations, all on a scale from 0 (*not at all*) to 10 (*very much*). These four items were integrated into an expectancy-violation score (average of the four items). The exact wording of and data for these four items are provided in the Supplemental Material available online.

Day 2: memory test. Seven days after Day 1, participants returned to the lab. Because both sleep (Diekelmann & Born, 2010) and the amount of rehearsal (Karpicke & Roediger, 2008) are known to affect memory formation, we asked participants to complete brief questionnaires on their sleep quality and duration between the 2 experimental days as well as on the amount of rehearsal in the aftermath of Day 1 (e.g., if and how often they talked about their experience with other people, if and how often they thought about the experience). Next, participants provided another saliva sample for later cortisol analysis and answered the mood questionnaire again. Thereafter, participants completed the critical recognition-memory test on a computer. This test included nine central items and 12 peripheral items that were part of the stress or control manipulation on Day 1 as well as 23 new distractor items that were categorized as central or peripheral on the basis of their similarity to the actually presented central or peripheral items. Distractor items were perceptually distinct but semantically related to the original stimuli. For instance, both original and distractor items included pictures, a sharpener, a system for blood pressure measurement, and a backpack. Trial order was randomized. On each trial, participants first saw a fixation cross with a variable duration of 1 s to 2 s, followed by an image of a previously seen or new item (Fig. 1). Participants were requested to indicate whether each item was old or new, as well as their confidence in that judgment, using the following options: "very certain old," "certain old," "rather old," "rather new," "certain new," or "very certain new." There was no time limit for the response, but participants were asked to respond quickly.

Statistical analysis

Subjective and physiological stress parameters were subjected to mixed-design ANOVAs with treatment (stress vs. control) and expectancy violation (prior information vs. no prior information) as between-subjects factors and time point of measurement as a within-subjects factor. For recognition memory, we focused our analyses on high-confidence hits (very certain, certain) as remembered central or peripheral items, in line with previous memory studies (Turk-Browne, Yi, & Chun, 2006; Wagner et al., 1998), because high-confidence hits are assumed to reflect actual recollection (Yonelinas, 1994). Analyses including all hits, irrespective of confidence, are provided in the Supplemental Material. To correct for potential biases in responding, we subtracted the individual false-alarm rate from the hit rate and focused on the resulting memory accuracy (Tambini, Rimmele, Phelps, & Davachi, 2017). Memory-accuracy scores were subjected to a mixed-design ANOVA with the between-subjects factors of treatment and expectancy violation and the within-subjects factor of item type (central vs. peripheral). Significant main or interaction effects were followed by post hoc tests that were Bonferroni-corrected (p_{corr}) if indicated. In case of violation of the sphericity assumption, Greenhouse-Geisser correction was applied. All reported p values are two-tailed. All statistical analyses were performed with SPSS Version 22. For the comparison of the informed and uninformed stress groups, we further performed a Bayesian analysis using JASP (Version 0.12; JASP Team, 2020). Here, we used a Cauchy distribution as the default prior with a spread parameter (r) of .707.

fNIRS recording and analysis

Cortical activation was measured with fNIRS during a 5-min baseline session and during the stress or control manipulation. We used a NIRScout System (NIRx Medical Technologies, Los Angeles, CA) with 16 sources and 16 detectors, forming 37 channels. The system was equipped with avalanche photooptodes, ensuring an optimal signal-to-noise ratio and short-distance detectors that acquire extracerebral hemodynamic signals that are regressed out from cerebral signals, thus allowing us to control for blood pressure differences between treatment conditions. On the basis of previous research on memory formation under stress and the salience network active during stressful events (Henckens et al., 2009; Hermans et al., 2014), we constructed our montage to cover, in addition to sensory and motor control areas, the following cortical regions: dorsolateral prefrontal cortex (dlPFC); frontal eye fields; inferior, middle, and superior temporal gyrus; temporal pole; TPJ; fusiform gyrus; and dorsal-posterior parietal cortex.

Data were preprocessed in nirsLAB (Version 2016.01; Xu, Graber, & Barbour, 2014). We identified the detector-saturation and interpolated-consecutive channels if required. Data quality of the channels was inspected, and if the coefficient of variation (CV) was greater than or equal to 15%, indicating a poor signal-to-noise-ratio, channels were excluded. In an individual, first-level analysis, we used prewhitening with autoregression and the regressors baseline and treatment, both modeled with a hemodynamic response function. Our analyses focused on oxygenated hemoglobin and the contrast treatment minus baseline. The preprocessed NIRS data were further processed using a MATLAB (MathWorks, Natick, MA) script that generated a matrix of the beta values across all channels and then integrated all channels that belonged to one topographical cluster, with each channel being weighed by the specificity of the channel for the respective brain region. The first-level contrast was then taken to a second, group level focusing on the interaction of Treatment × Expectancy

Table 1. Subjective and Cortisol Responses of Informed and Uninformed Participants to the Stress and Control Manipulations

| | Stress manipulation | | | | Control manipulation | | | |
| | Uninformed participants | | Informed participants | | Uninformed participants | | Informed participants | |
Variable	M	SD	M	SD	M	SD	M	SD
Stressfulness	7.17**	2.60	7.92**	2.28	3.61	2.19	3.39	1.85
Unpleasantness	8.25**	2.56	7.75**	2.45	3.35	1.80	3.09	2.07
Difficulty	8.08**	2.13	8.21**	1.74	2.87	1.39	3.35	2.29
Good vs. bad mood								
Before manipulation	34.50	4.38	35.42	3.15	34.70	3.27	33.83	5.37
After manipulation	26.83**	9.06	27.29**	6.56	34.09	4.38	32.43	5.34
Calmness vs. restlessness								
Before manipulation	32.54	5.96	34.58	3.69	33.04	4.42	33.30	4.90
After manipulation	25.50*	9.08	23.96**	6.97	30.78	5.79	32.09	5.38
Alertness vs. tiredness								
Before manipulation	28.46	6.26	30.42	6.58	29.91	5.92	29.17	5.00
After manipulation	27.54	6.55	26.33§	5.13	30.91	5.07	27.91	5.28
Salivary cortisol (nmol/l)								
Baseline	3.71	1.63	4.27	4.93	2.22	0.86	4.13	3.10
After 15 min	5.77	2.62	6.31§	3.51	2.48	1.35	8.18§	5.80
After 30 min	10.33§	6.46	11.59§	6.33	2.93	2.40	8.13	6.11
After 45 min	7.87§	5.17	9.05§§	4.81	2.23	1.71	7.64§	6.81
Baseline to peak	6.62§	6.56	7.31§	7.07	0.71	1.98	4.00	4.63

Note: Subjective assessments were given on a scale from 0 (*not at all*) to 10 (*very much*).
*$p < .05$, Bonferroni-corrected (vs. the informed control group). §$p < .05$, Bonferroni-corrected (vs. the uninformed control group). **$p < .01$, Bonferroni-corrected (vs. each of the control groups). §§$p < .01$, Bonferroni-corrected (vs. the uninformed control group).

Violation. The data of 14 participants (informed stress: $n = 4$, uninformed stress: $n = 6$, informed control: $n = 1$, uninformed control: $n = 3$) had to be removed because the exclusion of channels on the basis of the CV criterion resulted in the loss of three or more topographical clusters, thus leaving a sample of 80 participants for the fNIRS analysis.

Results

Successful stress induction

The subjective and cortisol data confirmed that the stress induction through the modified TSST was successful. As shown in Table 1, participants in the stress condition experienced the treatment as significantly more stressful, $F(1, 90) = 75.71$, $p < .001$, $\eta_p^2 = .46$; significantly more difficult, $F(1, 90) = 161.75$, $p < .001$, $\eta_p^2 = .64$; and significantly more unpleasant, $F(1, 90) = 106.50$, $p < .001$, $\eta_p^2 = .54$, than did participants in the control condition. Moreover, an ANOVA with group, prior information, and time point of measurement as factors revealed that negative mood, $F(1, 90) = 32.80$, $p < .001$, $\eta_p^2 = 0.27$; restlessness, $F(1, 90) = 32.13$, $p < .001$,

$\eta_p^2 = .26$; and wakefulness, $F(1, 90) = 5.40$, $p = .022$, $\eta_p^2 = .06$, increased from before to after the experimental treatment in the stress group but either did not increase or increased less in the control group (Table 1). In line with these subjective assessments, salivary cortisol levels increased significantly more strongly in response to the modified TSST than to the control manipulation, $F(1.95, 101.40) = 7.05$, $p = .001$, $\eta_p^2 = .12$. Cortisol concentrations were comparable at baseline, $t(55) = 1.01$, $p = .317$, $d = 0.27$, but significantly elevated in the stress group relative to the control group both 30 min after the treatment, $t(55) = 3.55$, $p_{corr} = .003$, $d = 0.94$, and 45 min after the treatment, $t(55) = 2.52$, $p_{corr} = .045$, $d = 0.67$ (Table 1). Accordingly, the baseline-to-peak difference, a single parameter reflecting the increase in cortisol, was significantly higher in the stress group than in the control group, $F(1, 53) = 10.52$, $p = .002$, $\eta_p^2 = .17$.

Importantly, the subjective and physiological response to the psychosocial stressor was not significantly modulated by the expectancy violation—Treatment × Expectancy Violation for the subjective ratings and baseline-to-peak increase in cortisol as well as Treatment × Expectancy Violation × Time Point of Measurement for the mood

questionnaire and cortisol: all Fs < 1.13, all ps > .328, all η_p^2s < .03; restlessness: $F(1, 90) = 3.42$, $p = .068$, $\eta_p^2 = .04$. As shown in Table 1, salivary cortisol levels were higher in informed than in uninformed participants immediately after the manipulation, which may be because of expectation effects and mistrust with respect to the announced control manipulation. Most importantly, however, the baseline-to-peak increase in salivary cortisol was comparable in informed and uninformed participants of the stress group, $t(25) = 0.26$, $p = .794$, $d = 0.10$. To explicitly test the evidence in favor of an absence of a difference between informed and uninformed participants of the stress groups, we further ran a Bayesian analysis for the baseline-to-peak increase in the informed versus uninformed stress group. This analysis revealed a Bayes factor (BF_{10}) of 0.368, which can be considered moderate to strong evidence for the null hypothesis that the informed and uninformed stress groups' cortisol responses to the stressor did not differ.

Prior information reduces expectancy violation

Participants' ratings of surprise and unexpectedness indicated that the detailed information before the treatment reduced, as predicted, the expectancy violation evoked by the treatment. Specifically, the integrated expectancy-violation score was significantly lower in informed participants than in uninformed participants, $F(1, 90) = 13.16$, $p < .001$, $\eta_p^2 = .20$. Moreover, the exposure to the stressful event was generally associated with a larger degree of expectancy violation than the control manipulation, $F(1, 90) = 22.15$, $p < .001$, $\eta_p^2 = .13$, reflecting the fact that the psychosocial stressor deviated more strongly from everyday experiences than the control manipulation. Accordingly, as shown in Figure 2, participants who underwent the psychosocial stressor with only minimal prior information reported the highest degree of expectancy violation (compared with each of the other three groups, all p_{corr}s ≤ .005), whereas participants who were well informed about the upcoming stressor reported a degree of expectancy violation comparable with the control groups' (both p_{corr}s ≥ .088).

Prior information reduces inferior temporal activity during a stressful event

During the stressful event and the control manipulation, we used fNIRS to record brain activity from cortical areas that have previously been implicated in memory
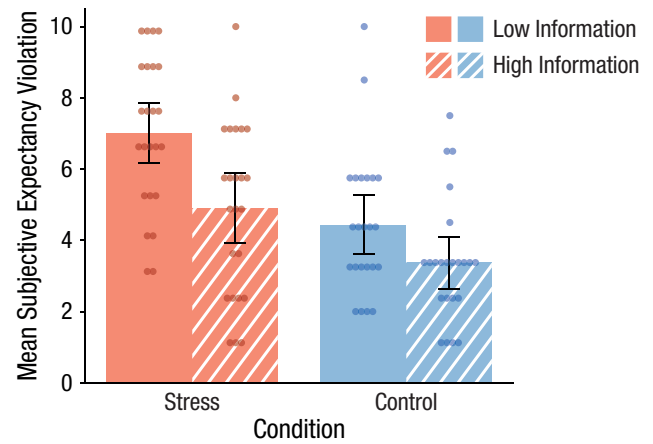

Fig. 2. Mean subjective expectancy violation in the stress and control groups, separately for participants given detailed and minimal information about the stressor. Data bars show means, and error bars show 95% confidence intervals. Dots represent individual participants' data.

formation and information processing under stress (Henckens et al., 2009; Hermans et al., 2014). Figure 3a shows striking differences in cortical activity between groups, with significant activity increases in the uninformed stress group that were not observed in any of the other three groups (i.e., activity was significantly reduced in stressed participants who received detailed information about the upcoming stressor). Accordingly, an overall analysis of the activity in these brain regions during the treatment (vs. baseline) revealed a significant Treatment × Expectancy Violation interaction, $F(1, 76) = 4.36$, $p = .040$, $\eta_p^2 = .05$, suggesting that the brain response to the stressful event was modulated by the expectancy violation. Follow-up analyses showed that the activity in the predefined cortical areas was higher in stressed participants who had received only minimal information about the treatment (i.e., a higher expectancy violation) than in stressed participants who were informed in detail about the upcoming stressor, $F(1, 36) = 4.72$, $p = .036$, $\eta_p^2 = .12$, whereas there was no effect of prior information in the control condition, $F(1, 40) = 0.74$, $p = .394$, $\eta_p^2 = .02$.

We hypothesized that the modulatory effect of expectancy violation on stress-related brain activity would not be equal across cortical areas but would be particularly pronounced in the inferior temporal cortex, an area involved in memory formation under stress. To test this hypothesis, we ran an additional analysis in which we directly compared the interactive effect of treatment and expectancy violation in four brain areas: (a) in the inferior temporal cortex as the candidate area for memory formation under stress (on the basis of previous neuroimaging evidence; Henckens et al.,
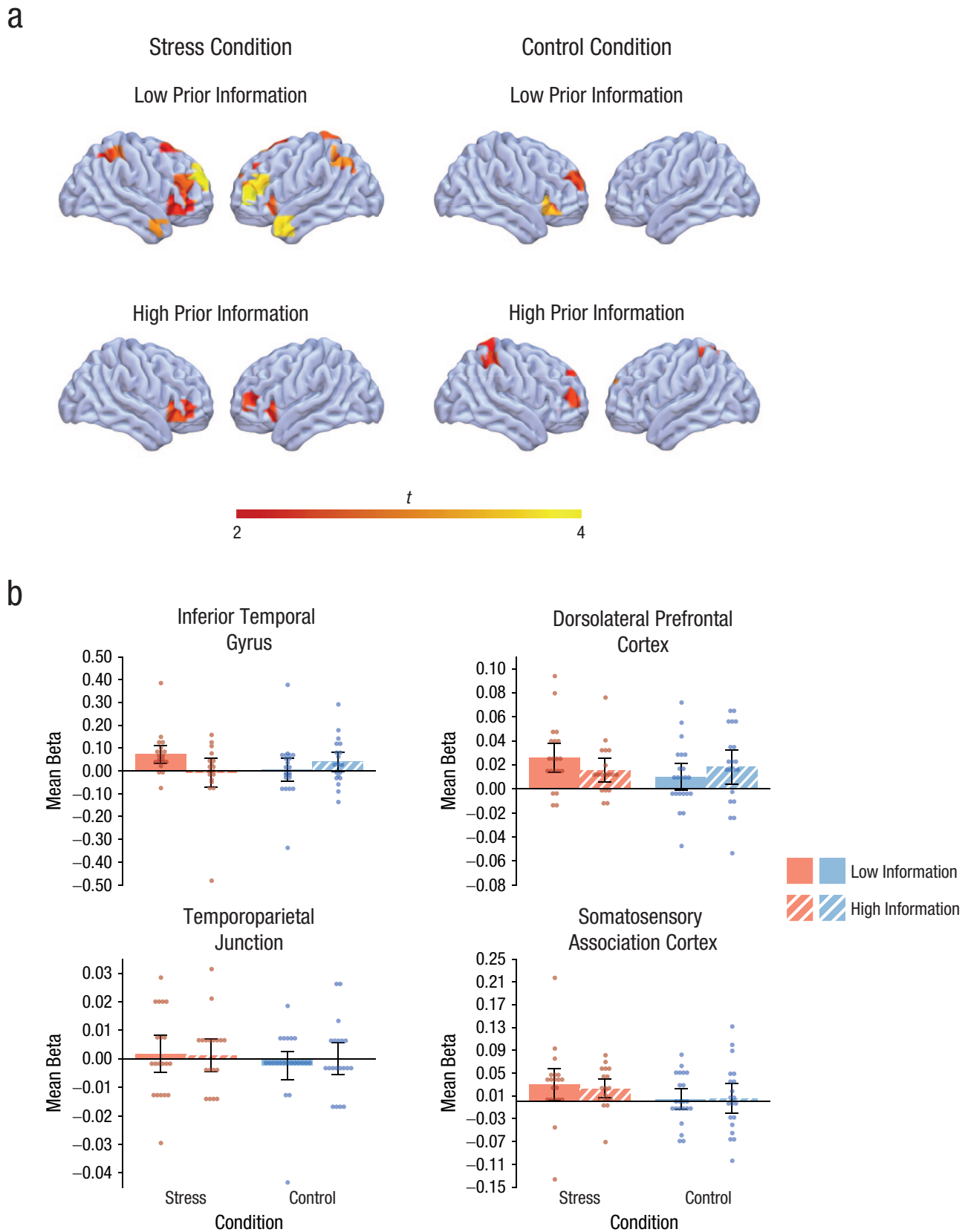
Fig. 3. Cortical activity during the treatment. The brain images (a) show activity in predefined cortical areas during the treatment (vs. baseline) as measured with functional near-infrared spectroscopy. Activity is shown separately for participants in the stress and control groups who were given detailed and minimal information about the stressor. The displayed activity has a threshold of $t = 2.00$, corresponding roughly to $p < .05$. The graphs (b) show beta coefficients for participants given detailed and minimal information in the stress and control groups, separately for the four regions of interest. Data bars show means, and error bars show 95% confidence intervals. Dots represent individual participants' data.

2009); (b) in the TPJ, which is also part of the salience network that may be preferentially recruited under stress (Hermans et al., 2014); (c) in the dlPFC, which is known to be highly sensitive to the influence of stress (Qin, Hermans, Van Marle, Luo, & Fernandez, 2009) but should be less involved in memory formation under stress; and (d) in the subcentral gyrus, an area that should be neither sensitive to the influence of stress nor involved in memory formation. This analysis yielded a significant Treatment × Expectancy Violation × Area interaction, $F(1.30, 99.06) = 3.72$, $p = .046$, $\eta_p^2 = .05$, indicating that the interactive influence of treatment and expectancy violation was indeed dependent on the specific brain area. Follow-up analyses showed that there was no Treatment × Expectancy Violation interaction in the subcentral gyrus, $F(1, 76) = 0.12$, $p = .726$, $\eta_p^2 < .01$; TPJ, $F(1, 76) = 0.27$, $p = .606$, $\eta_p^2 < .01$; or dlPFC, $F(1, 76) = 2.30$, $p = .13$, $\eta_p^2 = .03$. In the inferior temporal gyrus, however, there was a significant interaction of treatment and expectancy violation, $F(1, 76) = 5.11$, $p = .027$, $\eta_p^2 = .06$, showing that inferior temporal activity was significantly higher in stressed participants with only minimal prior information about the psychosocial stressor than in the informed stress group, $F(1, 36) = 4.74$, $p = .036$, $\eta_p^2 = .12$, whereas there was no influence of prior information in the control condition, $F(1, 40) = 0.97$, $p = .331$, $\eta_p^2 = .02$ (Fig. 3b).

Prior information attenuates memory boost for central features of a stressful event

Before the critical recognition test on Day 2 of the experiment, groups did not differ in wakefulness, restlessness, or salivary cortisol concentrations (all Fs < 2.02, all ps > .117, all η_p^2s < .06). Participants of the stress groups (vs. control groups) felt less positive on Day 2, $F(3, 89) = 3.27$, $p = .025$, $\eta_p^2 = .10$, but there were no differences between the informed and uninformed stress groups, $t(46) = -0.11$, $p = .915$, $d = 0.03$ (see Table S1 in the Supplemental Material). Moreover, groups did not differ in their sleep quality between Days 1 and 2 or the amount of rehearsal in the aftermath of Day 1 (all ps > .25; see Table S2 in the Supplemental Material).

Recognition performance, expressed as memory accuracy (Tambini et al., 2017), was analyzed by a Treatment × Expectancy Violation × Item Type (central vs. peripheral) ANOVA. This analysis revealed a significant three-way interaction, $F(1, 89) = 4.09$, $p = .046$, $\eta_p^2 = .04$, suggesting that the impact of stress on the memory for central versus peripheral information from the stress or control procedure depended critically on the expectancy violation associated with the treatment

(i.e., the amount of prior information). Follow-up analyses revealed a significant Treatment × Item Type interaction if there was only minimal information about the treatment, $F(1, 44) = 4.13$, $p = .048$, $\eta_p^2 = .04$, whereas this interaction effect was abolished by detailed prior information, $F(1, 45) = 0.57$, $p = .456$, $\eta_p^2 = .01$. As shown in Figure 4, stressed participants who received only minimal information about the stressor showed significantly better memory performance for central items than participants who received detailed information about the stressful event—central items: $t(46) = 2.14$, $p = .038$, $d = 0.62$; peripheral items: $t(46) = -0.68$, $p = .503$, $d = 0.19$—whereas there were no such differences in participants who underwent the control procedure (both ts < 1, both ps > .34, both ds < 0.29). Notably, in the uninformed stress group, memory performance was correlated with inferior temporal activity during the stressful event ($r = .52$, $p = .026$).

Follow-up analyses focusing on the hit rate and false-alarm rate, separately, revealed that the interactive effect of treatment and expectancy violation was mainly driven by the hit rate. More specifically, the hit rate for central elements of the stressful event was dependent on the degree of expectancy violation, as shown by a Treatment × Expectancy Violation ANOVA—central items: $F(1, 89) = 6.97$, $p = .010$, $\eta_p^2 = .07$; peripheral items: $F(1, 89) = 0.35$, $p = .557$, $\eta_p^2 < .01$—but there was no such interaction effect for the false-alarm rate—central items: $F(1, 89) = 0.83$, $p = .364$, $\eta_p^2 = .01$; peripheral items: $F(1, 89) = 5.42$, $p = .022$, $\eta_p^2 = .06$. There was a trend for a reduced false-alarm rate in the informed stress group compared with the informed control group, $F(1, 45) = 3.75$, $p = .059$, $\eta_p^2 = .08$. As shown in Table 2, the hit rate for central elements of the treatment was markedly increased—at least by 12%—in stressed participants with only minimal prior information compared with participants in each of the other three groups (all p_{corr}s < .033, all ds > 0.81).

As shown in Figure 4, the control groups' memory accuracy for central items was intermediate between the uninformed and informed stress groups' and did not significantly differ from the uninformed stress group's (both ts < 1.30, both p_{corr}s > .402, both ds < 0.39). Splitting the memory accuracy into the hit rate and false-alarm rate revealed that the uninformed stress group had a significantly higher hit rate for central elements of the episode than both the uninformed control group, $t(44) = 3.12$, $p_{corr} = .006$, $d = 0.92$, and the informed control group, $t(45) = 2.78$, $p_{corr} = .016$, $d = 0.81$ (Table 2). At the same time, the false-alarm rate for central items was higher in the uninformed stress group than in the uninformed control group, $t(44) = 2.20$, $p_{corr} = .066$, $d = 0.65$, and higher than in the informed control group, $t(45) = 2.21$, $p_{corr} = .064$, $d = 0.64$.
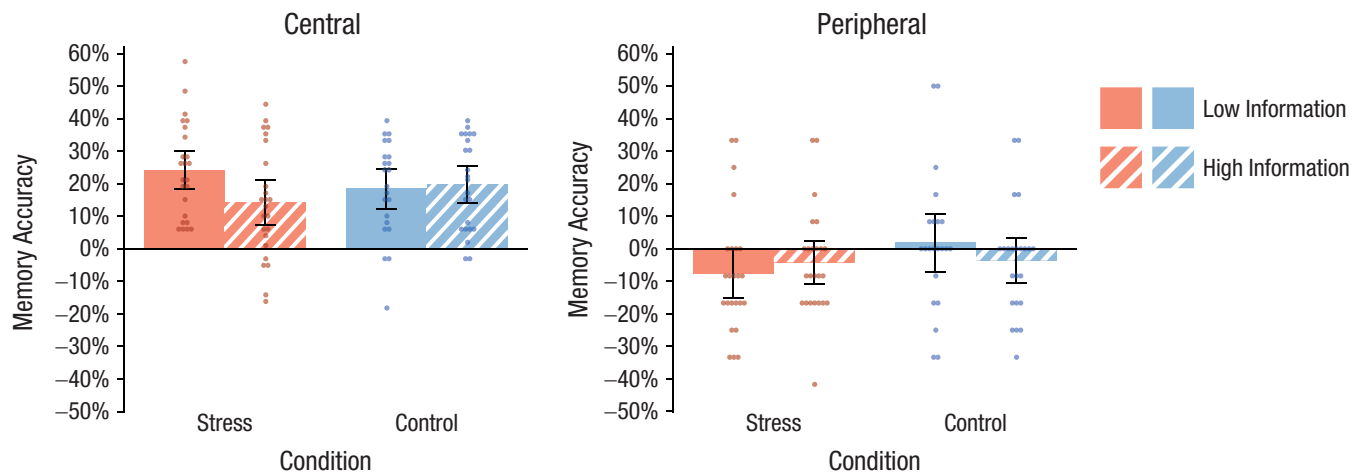
Fig. 4. Memory performance in the recognition test for central items (with which panel members had interacted and were therefore part of the stress manipulation) and peripheral items (which were not part of the stress manipulation). In each graph, results are shown separately for the participants in the stress and control groups who were given detailed and minimal information about the stressor. Data bars show means, and error bars show 95% confidence intervals. Dots represent individual participants' data.

No such differences were observed for the false-alarm rate for peripheral items (both $ts < 1.39$, both $p_{corr}s > .342$, both $ds < 0.41$), suggesting that the false-alarm rate was not generally elevated in the uninformed stress group but dependent on the content of the lures. Given that the lures were semantically related to the original items and thus involved a memory component, the present pattern of results suggests that, compared with memory in the nonstressed control participants, memory in the uninformed stress group was stronger but less specific, in line with known biases in emotional-recognition memory (White, Kapucu, Bruno, Rotello, & Ratcliff, 2014) and overgeneralized memory in PTSD (Schönfeld, Ehlers, Böllinghaus, & Rief, 2007).

Control variables

Overall, participants' levels of chronic stress, depressive mood, and trait anxiety were relatively low and comparable in the experimental groups (all $ps > .269$; Table S3 in the Supplemental Material).

Discussion

Stressful events are often vividly remembered. This memory boost for stressful events may be highly adaptive to survival but may contribute to PTSD and anxiety disorders. Whereas traditional accounts of emotional-memory enhancements emphasize the role of arousal-related hormones and neurotransmitters (Joëls et al., 2006; Schwabe et al., 2012), our recent account posits that the enhanced memory under stress may be driven by the expectancy violation evoked by stressful events (Trapp et al., 2018). Here, we tested a key assumption

of this alternate account, namely, that a reduction of the expectancy violation associated with a stressful encounter would abolish the memory enhancement. In line with this prediction, the present data show, for the first time, that detailed information about an upcoming psychosocial stressor prevents the enhanced memory for central elements of the stressful event. Using fNIRS to measure cortical activity during the stressful (or control) episode, we further link the enhanced memory formation under stress (without prior information) to the inferior temporal gyrus and show that the activity of this area under stress is critically modulated by the degree of expectancy violation associated with the stressful event.

Although expectancy violations (or prediction errors) have been most often studied in the context of associative learning (den Ouden, Friston, Daw, McIntosh, & Stephan, 2008; Rescorla & Wagner, 1972), recent research suggests that the role of expectancy violation extends to episodic memory. In particular, prediction errors during reward learning or Pavlovian fear conditioning have been shown to promote subsequent memory for trial-unique stimuli (Kalbe & Schwabe, 2020; Rouhani, Norman, & Niv, 2018). The present results dovetail with these findings but extend them in critical ways. First, we show an expectancy-violation-related memory enhancement for an aversive event using a rich, naturalistic episode rather than a sequence of unrelated and artificial stimuli (for further evidence showing an expectancy-violation-related enhancement of episodic memory per se for a naturalistic event, see Wahlheim & Zacks, 2019). Second, and even more importantly, we manipulated the degree of expectancy violation experimentally and demonstrated a causal role of the

Table 2. Hit (High Confidence) and False-Alarm Rates for Central (*n* = 9) and Peripheral (*n* = 12) Items in the Recognition Test on Day 2

| | Stress manipulation | | | | Control manipulation | | | |
| | Uninformed participants | | Informed participants | | Uninformed participants | | Informed participants | |
Item	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Central items								
Hit rate (high confidence)	53.70*	15.25	38.89	14.65	40.40	13.52	41.55	14.69
False-alarm rate	29.55	11.13	24.62	13.53	21.90	12.44	21.74	13.07
Peripheral items								
Hit rate (high confidence)	15.28	16.24	11.81	13.66	18.94	20.60	19.57	16.01
False-alarm rate	22.92	14.38	15.97	10.97	17.05	14.20	23.19	14.42

Note: The hit rate refers to the number of high-confidence hits relative to all old items.
*$p < .05$, Bonferroni-corrected (vs. each of the other groups).

available information—and hence the extent of expectancy violation—in the enhanced memory formation for the stressful event. Moreover, our data show that stress, and presumably the related expectancy violation, does not promote memory for all elements of the stressful episode equally. Instead, stress specifically enhanced the memory for central elements of the event, in line with previous research (Kensinger et al., 2007; Wiemers et al., 2013), and this specific memory boost was blocked by providing detailed information about the upcoming stressor. Previous research has suggested that the memory enhancement for central elements would come at the cost of memory for peripheral details (Kensinger et al., 2007). Here, we did not observe a stress-induced impairment in memory for peripheral stimuli, which may be because of the overall rather moderate memory performance for peripheral details that might be explained by the incidental encoding, in combination with the retention interval of 1 week.

One might argue that the detailed information about the stressful event may have reduced the physiological (and subjective) stress response and that this attenuated stress response then led to the reduced memory formation under stress. Importantly, however, the stress responses of informed and uninformed participants did not differ. Both the subjective and the cortisol response to the psychosocial stressor were largely unaffected by the amount of prior information and, by implication, expectancy violation. Although a Bayesian analysis provided strong evidence that prior information did not change the cortisol response, the critical comparison of the informed and uninformed stress groups in terms of the cortisol response relied on a rather small sample because of the loss of saliva samples for some of the participants, which is a caveat of the present study. Although we cannot rule out the possibility that informed and uninformed participants differed in other

stress parameters than the ones measured here, it is remarkable that the subjective and cortisol responses were comparable in both of these groups and that the mere availability of detailed information about the stressor attenuated the memory for the stressful event. This finding suggests that stress hormones, which are well known to enhance memory formation regardless of the cognitive processes engaged during learning (de Quervain et al., 2017; Sandi, 2011), and expectancy violation might be two separate routes to enhanced memory formation under stress.

The expectancy violation may exert its effect on memory formation through altered attentional, perceptual, or motivational processing, all of which is known to be driven by prediction errors (den Ouden, Kok, & de Lange, 2012). A closely related potential mechanism through which an expectancy violation may affect memory formation under stress is an altered recruitment of the salience network. Our fNIRS data, however, showed that the activity of the TPJ—an integral part of the salience network—under stress was not modulated by prior information, suggesting that the degree of expectancy violation did not modulate the salience network as a whole. Instead, we obtained an expectancy-violation-related decrease in activity under stress specifically in the inferior temporal gyrus. The inferior temporal gyrus may be a prime candidate for expectancy-violation-dependent memory formation under stress, given that it is a hub for the integration of mnemonic, attentional, and visual processing (Miyashita, 1993). Furthermore, our present findings on the role of the inferior temporal gyrus are consistent with previous fMRI evidence pointing to this area as a hot spot for memory formation under stress (Henckens et al., 2009). Nevertheless, the conclusion that the inferior temporal gyrus is the key region mediating the impact of expectancy violation on memory formation for stressful

events might be premature because fNIRS measures activity from predefined cortical regions but not from subcortical areas such as the amygdala or hippocampus, which are known to play a prominent role in emotional memory (Joëls et al., 2006; Schwabe et al., 2012). However, our data show that the manipulation of the expectancy violation evoked by the stressful event is directly linked to altered activity in a brain area highly relevant for memory formation under stress.

In conclusion, our findings demonstrate, for the first time, a critical role of prior information about an upcoming stressful event and, by implication, the extent of expectancy violation evoked by this event for the later memory of the stressful encounter, thus supporting a recent cognitive account of the superior memory formation for stressful events (Trapp et al., 2018). Although the traditional arousal view of memory formation and the proposed expectancy-related view are not mutually exclusive and there is evidence for an overlap between arousal and predictability (de Berker et al., 2015), the view held may guide the approach taken to interfere with overly strong emotional memory. The arousal model inspired mainly pharmacological attempts to modify emotional-memory formation by interfering with noradrenergic or glucocorticoid signaling (de Quervain et al., 2017; Pitman et al., 2012). The more cognitive, expectancy-violation model, in turn, points to psychological interventions focused on a cognitive preparation for a potential stressor. We acknowledge that many stressors are characterized by their unpredictability, making a cognitive preparation difficult. However, in some contexts—for instance, during military actions or in emergency settings—stressful events are more predictable and more clearly defined. Our findings suggest that in such contexts, detailed information and extensive training may be essential to prevent the emergence of debilitating memories.

Transparency

ORCID iD

Lars Schwabe https://orcid.org/0000-0003-4429-4373

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797620958650

References

Beck, A. T., & Steer, R. A. (1987). *Beck Depression Inventory manual*. San Antonio, TX: Psychological Corp.

de Berker, A. O., Rutledge, R. B., Mathys, C., Marshall, L., Cross, G. F., Dolan, R. J., & Bestmann, S. (2015). Computations of uncertainty mediate acute stress responses in humans. *Nature Communications*, *7*, Article 10996. doi:10.1038/ncomms10996

den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2008). A dual role for prediction error in associative learning. *Cerebral Cortex*, *19*, 1175–1185.

den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*, Article 548. doi:10.3389/fpsyg.2012.00548

de Quervain, D., Schwabe, L., & Roozendaal, B. (2017). Stress, glucocorticoids and memory: Implications for treating fear-related disorders. *Nature Reviews Neuroscience*, *18*, 7–19.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*, 114–126.

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*, 175–191.

Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction error drive one-shot declarative learning? *Journal of Memory and Language*, *94*, 149–165.

Henckens, M. J. A. G., Hermans, E. J., Pu, Z., Joëls, M., & Fernandez, G. (2009). Stressed memories: How acute stress affects memory formation in humans. *The Journal of Neuroscience*, *29*, 10111–10119.

Hermans, E. J., Henckens, M. J. A. G., Joëls, M., & Fernandez, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends in Neurosciences*, *37*, 304–314.

Jasp Team. (2020). JASP (Version 0.12) [Computer software]. Retrieved from https://jasp-stats.org/

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*, 152–158.

Kalbe, F., & Schwabe, L. (2020). Beyond arousal: Prediction error related to aversive events promotes episodic memory formation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*, 234–246.

Karpicke, J. D., & Roediger, H. L., III (2008). The critical importance of retrieval for learning. *Science*, *319*, 966–968.

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2007). Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. *Journal of Memory and Language*, *56*, 575–591.

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*, 76–81.

Miyashita, Y. (1993). Inferior temporal cortex: Where visual perception meets memory. *Annual Review of Neuroscience*, *16*, 245–263.

Pitman, R. K., Rasmusson, A. M., Koenen, K. C., Shin, L. M., Orr, S. P., Gilbertson, M. W., . . . Liberzon, I. (2012). Biological studies of post-traumatic stress disorder. *Nature Reviews Neuroscience*, *13*, 769–787.

Qin, S., Hermans, E. J., Van Marle, H. J. F., Luo, J., & Fernandez, G. (2009). Acute psychosocial stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biological Psychiatry*, *66*, 25–32.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokassy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Rouhani, N., Norman, K., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1430–1443.

Sandi, C. (2011). Glucocorticoids act on glutamatergic pathways to affect memory processes. *Trends in Neurosciences*, *34*, 165–176.

Sandi, C., Loscertales, M., & Guaza, C. (1997). Experience-dependent facilitating effect of corticosterone on spatial memory formation in the water maze. *European Journal of Neuroscience*, *9*, 637–642.

Schönfeld, S., Ehlers, A., Böllinghaus, I., & Rief, W. (2007). Overgeneral memory and suppression of trauma memories in post-traumatic stress disorder. *Memory*, *15*, 339–352.

Schulz, P., Schlotz, W., & Becker, P. (2004). *Trier Inventory for Chronic Stress*. Gottingen, Germany: Hogrefe.

Schwabe, L., Bohringer, A., Chatterjee, M., & Schachinger, H. (2008). Effects of pre-learning stress on memory for neutral, positive and negative words: Different roles of cortisol and autonomic arousal. *Neurobiology of Learning and Memory*, *90*, 44–53.

Schwabe, L., Joëls, M., Roozendaal, B., Wolf, O. T., & Oitzl, M. S. (2012). Stress effects on memory: An update and integration. *Neuroscience and Biobehavioral Reviews*, *36*, 1740–1749.

Sevenster, D., Beckers, T., & Kindt, M. (2013). Prediction error governs pharmacologically induced amnesia for learned fear. *Science*, *339*, 830–833.

Sinclair, A. H., & Barense, M. D. (2019). Prediction error and memory reactivation: How incomplete reminders drive reconsolidation. *Trends in Neurosciences*, *42*, 727–739.

Spielberger, C. D., Gorsuch, R. L., & Luchene, R. E. (1970). *The State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1994). Testtheoretische Analysen des Mehrdimensionalen Befindlichkeitsfragebogens (MDBF) [Test theoretical analyses of the Multidimensional Mood Questionnaire (MDBF)]. *Diagnostica*, *40*, 320–328.

Tabachnik, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston, MA: Allyn & Bacon.

Tambini, A., Rimmele, U., Phelps, E. A., & Davachi, L. (2017). Emotional brain states carry over and enhance future memory formation. *Nature Neuroscience*, *20*, 271–278.

Trapp, S., O'Doherty, J. P., & Schwabe, L. (2018). Stressful events as teaching signals for the brain. *Trends in Cognitive Sciences*, *22*, 475–478.

Turk-Browne, N. B., Yi, D. J., & Chun, M. M. (2006). Linking implicit and explicit memory: Common encoding factors and shared representations. *Neuron*, *49*, 917–927.

Vogel, S., & Schwabe, L. (2016). Stress in the zoo: Tracking the impact of stress on memory formation over time. *Psychoneuroendocrinology*, *71*, 64–72.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., . . . Buckner, R. L. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, *281*, 1188–1191.

Wahlheim, C. N., & Zacks, J. M. (2019). Memory guides the processing of event changes for older and younger adults. *Journal of Experimental Psychology: General*, *148*, 30–50.

White, C. N., Kapucu, A., Bruno, D., Rotello, C. M., & Ratcliff, R. (2014). Memory bias for negative emotional words in recognition memory is driven by effects of category membership. *Cognition and Emotion*, *28*, 867–880.

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*, 2268–2277.

Xu, Y., Graber, H. L., & Barbour, R. L. (2014, April 26–30). *nirsLAB: A computing environment for fNIRS neuroimaging data analysis*. Paper presented at the 2014 Biomedical Optics Meeting of The Optical Society, Miami, Florida.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1341–1354.