Neurobiology of Learning and Memory 134 (2016) 392–399

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

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

The impact of psychosocial stress on conceptual knowledge retrieval

Christian J. Merz [a,b,*], Florian Dietsch [a,c], Michael Schneider [c]

[a] Department of Biological and Clinical Psychology, University of Trier, Johanniterufer 15, 54290 Trier, Germany
[b] Institute of Cognitive Neuroscience, Department of Cognitive Psychology, Ruhr-University Bochum, Universitätsstr. 150, 44780 Bochum, Germany
[c] Department of Educational Psychology, University of Trier, Universitätsring 15, 54286 Trier, Germany

ARTICLE INFO

Article history:
Received 13 May 2016
Revised 21 July 2016
Accepted 28 August 2016
Available online 30 August 2016

Keywords:
Cortisol
Glucocorticoids
SECPT
Naïve theory
Scientific theory
Stress hormones

ABSTRACT

The acquisition of conceptual knowledge in scientific domains is among the central aims of school instruction because this semantic declarative knowledge helps individuals make interferences and explain complex phenomena. Recent research shows that naïve concepts acquired during childhood persist in long-term memory long after learning the scientifically correct concepts in school. In this study, we investigated the effects of stress on the retrieval of these conceptual representations. To this end, 40 healthy men were randomly assigned to either psychosocial stress or a control condition and evaluated, as quickly and accurately as possible, statements that were compatible with scientific concepts or incompatible with those concepts. Some of these statements were true and some were false. Incompatible statements in this case are statements which are in line with adults' scientific concepts, but not with children's naïve theories. In contrast, compatible statements are in line with both. Stress induction was successful as evidenced by increases in blood pressure and cortisol concentrations in the stress group compared to the control group. Responses were delayed and less accurate for incompatible compared to compatible statements. Psychosocial stress had no main effect on retrieval, but abolished reaction time differences on false- vs. true-incompatible statements. This effect was mirrored in correlations between individuals' cortisol increases and reaction times. These results suggest that stress, as embodied by increases in cortisol concentrations, interferes with the retrieval of conceptual knowledge. They help to better understand conceptual knowledge retrieval in real-life situations such as examinations or problem solving in the workplace.

1. Introduction

A student completing an important exam, a pilot trying to fix a problem with the airplane mid-air, and a scientist interpreting data from an experiment under time pressure all have one thing in common: they need to retrieve conceptual knowledge, that is, semantic declarative memories about scientific concepts and the world, from their long-term memory while at the same time experiencing stress. Despite the general importance of the topic, little is known about the effects of stress on the retrieval of conceptual knowledge from memory.

It is widely acknowledged that stress influences memory retrieval processes in general through the release of stress mediators such as (nor)epinephrine after activation of the sympathetic nervous system (SNS) and glucocorticoids (GCs; cortisol in humans) resulting from stimulation of the hypothalamic-pituitary-adreno cortical (HPA) axis (Schwabe, Joëls, Roozendaal, Wolf, & Oitzl, 2012; Wolf, 2009). GCs can cross the blood-brain-barrier and bind to mineralocorticoid receptors and glucocorticoid receptors (GRs), both located in limbic regions such as the hippocampus, although GRs are also and predominantly expressed in prefrontal areas (Perlman, Webster, Herman, Kleinman, & Weickert, 2007). Both receptor types play a crucial role in memory functions in that stress can exert beneficial or impairing effects depending on the timing of stress onset (Het, Ramlow, & Wolf, 2005; Schwabe & Wolf, 2013; Schwabe, Wolf, & Oitzl, 2010). After exposure to stress or after GC administration, impairments can be observed in working memory (Oei, Everaerd, Elzinga, van Well, & Bermond, 2006; Schoofs, Preuβ, & Wolf, 2008; Schoofs, Wolf, & Smeets, 2009) as well as in the retrieval of declarative (De Quervain, Roozendaal, Nitsch, McGaugh, & Hock, 2000; Kuhlmann, Piel, & Wolf, 2005), social (Merz, Wolf, & Hennig, 2010; Takahashi et al., 2004) and autobiographical memories (Buss, Wolf, Witt, & Hellhammer, 2004; Schlosser et al., 2010). Neuroimaging studies have revealed

* Corresponding author at: Institute of Cognitive Neuroscience, Department of Cognitive Psychology, Ruhr-University Bochum, Universitätsstr. 150, 44780 Bochum, Germany.

E-mail addresses: Christian.J.Merz@rub.de (C.J. Merz), s1fldiet@uni-trier.de (F. Dietsch), mschneid@uni-trier.de (M. Schneider).

http://dx.doi.org/10.1016/j.nlm.2016.08.020
1074-7427/

that the hippocampus and the prefrontal cortex are involved in these stress effects on memory retrieval (Oei et al., 2007; Schwabe & Wolf, 2013).

Conceptual knowledge consists of general knowledge about everyday or scientific concepts and their interrelations (Carey, 2009), encoded in semantic declarative memory (Yee, Chrysikou, & Thompson-Schill, 2013). Individuals can use conceptual knowledge to categorize objects, to understand principles, to generalize, and to make predictions (Machery, 2010). Conceptual knowledge is used on a task-specific basis as a cognitive tool to construct mental models in working memory (Schnotz & Preuβ, 1997; Vosniadou & Brewer, 1992).

In many domains, children first acquire so-called naïve concepts in everyday life (e.g., the sun orbits the earth) and only later learn the scientifically correct concepts in school (cf. Carey, 1992). Several theories of conceptual change imply that newly learned scientific concepts overwrite and thus replace earlier naïve concepts (e.g. Carey, 2009; Vosniadou, 1994). Contrary to this classical view, recent empirical studies found that naïve concepts and scientific concepts coexisted in learners' long-term memory even after school instruction (Alibali & Sidney, 2015; Kelemen, Rottman, & Seston, 2013; Schneider & Hardy, 2013). The concepts are activated depending on the situational context, so that learners are not necessarily aware of the inconsistencies in their knowledge base (diSessa, Gillespie, & Esterly, 2004).

Recent reaction time studies demonstrate that naïve and scientific concepts do not only coexist in memory but actually interfere with each other during retrieval under time pressure (Babai, Sekal, & Stavy, 2010; Kelemen et al., 2013; Potvin, Masson, Lafortune, & Cyr, 2015; Shtulman & Harrington, 2016). Whereas most of these studies tested for knowledge in a single content domain only, Shtulman and Valcarcel (2012) provided consistent evidence from 50 concepts in ten scientific content domains (astronomy, evolution, fractions, genetics, germs, matter, mechanics, physiology, thermodynamics and waves). For each concept, they constructed four statements. One item conflicted with the naïve concept as well as with the scientific concept (compatible-false, e.g. "rain produces heat"); one item was in line with the naïve concept as well as with the scientific concept (compatible-true, e.g. "ovens produce heat"); one item was in line with the naïve concept but conflicted with the scientific concept (incompatible-false, e.g. "coats produce heat"); and one item conflicted with the naïve concept but was in line with the scientific concept (incompatible-true, e.g. "pressure produces heat"). The participants evaluated the truth value of each statement as quickly and accurately as possible by pressing one of two buttons. In all ten domains, participants consistently showed less accurate and slower responses when verifying incompatible compared to compatible statements. This is explained by the fact that interferences between naïve and scientific concepts need to be resolved in working memory for incompatible but not for compatible statements.

Shtulman and Valcarcel (2012) also found reaction time differences between scientifically true and scientifically false statements. This is a typical finding in sentence verification tasks (Neubauer & Freudenthaler, 1994). Recent brain imaging data (Marques, Canessa, & Cappa, 2009) suggest that the verification of true statements activates the left inferior parietal cortex and the caudate nucleus which are both involved in matching processes. In contrast, the verification of false statements involves the frontopolar cortex and might involve the evaluation of contradictions.

Taken together, research shows that naïve concepts persist after formal instruction on a topic and can interfere with later acquired scientific concepts, suggesting that scientific knowledge inhibits rather than replaces naïve knowledge. Since these complex causal thinking tasks involve increased activation of prefrontal regions (Brault Foisy, Potvin, Riopel, & Masson, 2015), stress might impair conceptual knowledge retrieval via its impact on prefrontal functioning (Oei et al., 2007; Schwabe & Wolf, 2013). The strengths of these effects might differ between the evaluations of compatible versus incompatible statements (Shtulman & Valcarcel, 2012) and true versus false statements (Marques et al., 2009) because these processes rely on partly different brain regions and cognitive resources.

The objective of this study was to examine the consequences of acute psychosocial stress and the associated endocrine response of the SNS and the HPA axis on conceptual knowledge retrieval. We hypothesized that stress would reduce conceptual knowledge retrieval in terms of a slower reaction time and a lower accuracy of responses, especially for false-incompatible statements. These false-incompatible statements require more conflict monitoring and interference resolution than the other three types of statements involving the recruitment of prefrontal areas (Marques et al., 2009), which are especially prone to stress effects due to the predominant expression of GRs in these areas (Perlman et al., 2007).

2. Material and methods

2.1. Participants

Participants were recruited through email announcements at the University of Trier, Germany, or by personally addressing them. Inclusion criteria comprised an age between 18 and 35 years and a body-mass-index between 18 and 27 kg/m^2. Due to known sex differences in cortisol responses to stress and in stress-mediated memory effects (Kudielka, Hellhammer, & Wüst, 2009; Merz & Wolf, 2015) only men were included in the present sample. Exclusion criteria were regular medication intake, any history of psychiatric treatment or a current psychiatric (e.g., depression or anxiety disorder) and/or somatic disease (e.g., high blood pressure, Raynaud's disease or allergies), especially endocrine diseases known to influence endogenous hormone levels (e.g., hyper-/hypothyroidism), and smoking. The study was approved by the local ethics committee of the University of Trier. Written informed consent was provided prior to participation in the study.

2.2. Procedure

Experimental sessions were run between 1 and 6 p.m. and participants had to have been awake for at least 3 h before testing in order to control for circadian fluctuations in salivary cortisol (Kudielka et al., 2009). They were instructed to refrain from intense physical exercise, eating and drinking anything but water for at least 90 min before the experiment.

On arrival, participants were randomly assigned to one of two groups comprising 20 persons each (socially evaluated cold-pressor test (SECPT) vs. warm water control condition; Schwabe, Haddad, & Schächinger, 2008). In both conditions, participants had to immerse their dominant hand up to the elbow into water, with a temperature between 0 and 3 °C in the SECPT and between 36 and 37 °C in the control condition. A neutral female experimenter only present for the duration of the SECPT videotaped and observed participants during the SECPT, while the experimenter executing the rest of the study stayed in the background. In the warm water control condition, neither videotaping nor observation took place. In both conditions, participants were instructed to remove their arm from the water after 3 min. If they did not manage to keep their arms in the ice water in the SECPT for this duration, they were instructed to hold their hands above the water for the remaining time (only one participant took his hand out of the cold water after 70 s).

The experimenter started the testing session by explaining cognitive performance testing, saliva sampling, and the recording of heart rate and blood pressure to the participants, as well as informing them that they would either be undergoing a challenging situation or a control condition. Following this, participants answered a demographic questionnaire asking for information on their final high school grades in mathematics, physics, chemistry, biology, German and the overall grade of their high school diploma. They were then instructed to give the first saliva sample. After a resting phase of 10 min, blood pressure and heart rate were measured. Afterwards, participants underwent the experimental manipulation (stress vs. control condition) with blood pressure and heart rate recorded simultaneously. After cessation, participants answered three questions concerning their subjective appraisal of the task (cf. 2.4). The second saliva sample was then provided and blood pressure and heart rate were measured, followed by a 10 min recovery period in which participants answered questionnaires as a filler task. Twenty minutes after stress onset, participants were asked for a third saliva sample, followed by a computer-based task adapted from Shtulman and Valcarcel (2012). After completion, participants provided the fourth saliva sample. Finally, they were debriefed and received either partial course credits or €10 as a monetary compensation.

2.3. Assessment of conceptual knowledge

Stimuli were presented with E-Prime® 2.0 (Psychology Software Tools, Sharpsburg, PA, USA) on a screen with a distance of 70 cm to the participants. The stimulus material was adapted from the study by Shtulman and Valcarcel (2012). The data were collected in Germany. Accordingly, the test items were translated into German by two native speakers of German independently. Discrepancies were resolved by discussion. To minimize the issue of changing stress parameters during a long knowledge test, we limited our material to 20 of the original 50 concepts, that is, two concepts from each of the ten content domains. For each of the 20 concepts, we chose a true-compatible (e.g. "ovens produce heat"), a false-compatible (e.g. "rain produces heat"), a true-incompatible (e.g. "pressure produces heat"), and a false-incompatible statement (e.g. "coats produce heat"), resulting in a total of 80 statements. Participants had to decide whether or not the presented statement was scientifically correct by pushing one of two labeled keys on the keyboard. The evaluation of the 80 statements took about 9 min on average.

The stimulus order was randomized for each person. Statements were preceded by a fixation cross presented for 500 ms and the time between the offset of a statement and the onset of the next one was 1500 ms. Differing from the original procedure, we did not match statements in blocks per domain or define a time limit for statement verification. Results from our own prior research have shown that the task induces reliable and consistent effects on reaction time and accuracy interference. In order to analyze the performance, reaction time and accuracy were recorded.

2.4. Measurement and analysis of the stress response

In order to verify activation of the SNS, heart rate as well as systolic and diastolic blood pressure were measured using an automatic upper arm blood pressure monitor (Bosch+Sohn, Jungingen, Germany). The cuff was placed 2 cm above the elbow of the non-dominant arm. In order to avoid measurement errors, participants were instructed to avoid speaking and moving during the procedure. Measurements took place 7 min before stress onset (baseline), during the 3 min of the stressor (peak) and 5 min after offset of the stressor (post). Assessment was carried out three times in a row at each time point in order to calculate the mean values of heart rate and blood pressure within a time window of 3 min.

In order to assess activation of the HPA axis, we collected saliva samples using Salivette collection devices (Sarstedt, Nuembrecht, Germany) 20 min before (baseline) as well as immediately (+3 min), 20 and 33 min after onset of the experimental procedure. All samples were stored at −20 °C until analysis. The fraction of free unbound salivary cortisol was determined using a Dissociation-Enhanced Lanthanide Fluorescent Immunoassay as described previously by Dressendörfer, Kirschbaum, Rohde, Stahl, and Strasburger (1992). The limit of detection was 0.5 nmol/l for saliva cortisol. Inter- and intra-assay coefficients of variance were below 9.0% and 6.7%, respectively.

Immediately after cessation of the stress or control procedure, participants rated how stressful, painful, and unpleasant they had experienced the respective procedure to be on a scale ranging from 0 ("not at all") to 100 ("very much"; ratings adopted from Schwabe et al., 2008).

2.5. Statistical analyses

Data were analyzed using the SPSS 20.0 software (SPSS Inc., Chicago, USA). Demographic variables and ratings of stress appraisal were analyzed using independent *t*-tests between the stress and the control group. Repeated measures analyses of variance (ANOVA) were applied to investigate changes in physiological parameters with the repeated measurement factor time (baseline, during and post hand immersion for heart rate and systolic and diastolic blood pressure; baseline, +3, +20 and +33 min for changes in cortisol concentrations) as well as the between-groups factor group (stress vs. control).

To investigate accuracy in the cognitive task, we calculated the mean accuracy value of classifications regarding each of the four statement types (true-compatible/false-compatible/true-incompatible/false-incompatible) for each participant. In the statistical analysis of reaction times, we only included reaction times stemming from statements answered correctly. Median reaction time was computed per statement category for each participant. The influence of stress on accuracy and reaction times was analyzed running separate ANOVA with the within-subjects factors scientific correctness (true vs. false) and compatibility (compatible vs. incompatible) and the between-subjects factor group.

Pearson product-moment correlations were performed to test whether the physiological stress response is associated with reaction times and/or accuracy during conceptual knowledge retrieval. Increases in heart rate (Δ heart rate), blood pressure (Δ sys, Δ dia) and cortisol (Δ cort) were calculated by subtracting the peak value of the respective stress marker (during hand immersion for heart rate and blood pressure; +20 min for cortisol) from the baseline value. Additionally, the area under the curve with respect to increase (AUC_i; Pruessner, Kirschbaum, Meinlschmid, & Hellhammer, 2003) was computed for cortisol and correlated with reaction times and accuracy variables. Partial correlation analyses were run on the same set of variables to control for the factor group.

Where appropriate, Greenhouse-Geisser degrees of freedom adjustment in case of a violation of sphericity was applied and effect sizes (partial η^2) are reported accordingly. A value of $p < 0.05$ was accepted to mark statistical significance.

3. Results

3.1. Demographic data

Students' grade point average from their high school diploma (Abitur: possible range: 1.0–4.0) had a sample mean of 2.3 (*SD* = 0.7; range from 1.2 to 3.9) indicating a broad range of competence in our sample. In addition, subject of study ranged from psychology, business administration and law, bio-/geoscience, education over political science, economic sociology, computer science, civil and industrial engineering to food economics, history, Japanology and energy management, again indicating a broad range of knowledge in the present sample. Participants in the control group were slightly older compared to participants in the stress group ($t_{(30.2)} = 2.85$, $p = 0.008$, $\eta^2 = 0.21$). All results reported below concerning differences between the stress and the control group remain statistically significant when age is added as covariate. No group differences were observed regarding BMI or high school grades (all $p > 0.05$: cf. Table 1).

3.2. Stress response

Higher cortisol concentrations were found in the stress compared to the control group over the course of time (time × group interaction; $F_{(1.5,56.9)} = 12.16$, $p < 0.001$, $\eta^2 = 0.24$). Follow-up analysis affirmed significant between-groups differences 20 min ($t_{(28.4)} = 3.48$, $p = 0.002$, $\eta^2 = 0.30$) and 33 min ($t_{(38)} = 2.42$, $p = 0.020$, $\eta^2 = 0.13$) after, but not before stress onset (both $t < 1.73$, both $p > 0.05$; cf. Fig. 1). No significant differences were obtained for the main effects of time and group (both $F < 3.12$, both $p > 0.05$).

For heart rate, a significant time × group interaction was identified ($F_{(1.4,54.7)} = 7.96$, $p < 0.003$, $\eta^2 = 0.17$). Although no significant between-groups differences occurred at the three sampling points (all $t < 1.52$, all $p > 0.13$), the stress group showed a significant increase from baseline to the sampling point during hand immersion ($t_{(19)} = 2.48$, $p = 0.023$, $\eta^2 = 0.24$) and a significant decrease from the sampling point during to that post hand immersion ($t_{(19)} = 2.59$, $p = 0.018$, $\eta^2 = 0.26$; cf. Table 1). No such effects were observed in the control group (all $t < 1.39$, all $p > 0.05$).

Table 1
Mean (±SEM) age, body-mass-index, high school diploma (grade point average; GPA), heart rate and blood pressure data as well as subjective ratings separated for the stress and control group.

	Control group	Stress group
Demographics		
Age (years)	25.1 ± 3.3	22.7 ± 1.9
Body-mass-index	24.3 ± 1.9	23.1 ± 2.2
High school diploma GPA	2.3 ± 0.2	2.4 ± 0.2
Systolic blood pressure (mmHg)		
Baseline	126.0 ± 10.4	130.9 ± 11.2
During hand immersion	127.0 ± 10.6	157.2 ± 16.0
5 min after stress/control	121.0 ± 9.4	127.0 ± 9.3
Diastolic blood pressure (mmHg)		
Baseline	75.9 ± 9.6	77.6 ± 7.4
During hand immersion	76.9 ± 10.0	98.3 ± 8.3
5 min after stress/control	75.2 ± 12.9	80.2 ± 7.9
Heart rate (beats per minute)		
Baseline	63.7 ± 8.1	60.4 ± 8.7
During hand immersion	62.0 ± 9.6	67.5 ± 12.9
5 min after stress/control	64.4 ± 9.5	59.8 ± 9.8
Subjective ratings after stress/control condition		
Stressful	3.6 ± 6.7	41.0 ± 32.9
Painful	1.0 ± 4.5	49.5 ± 29.5
Unpleasant	4.1 ± 6.0	48.0 ± 25.9

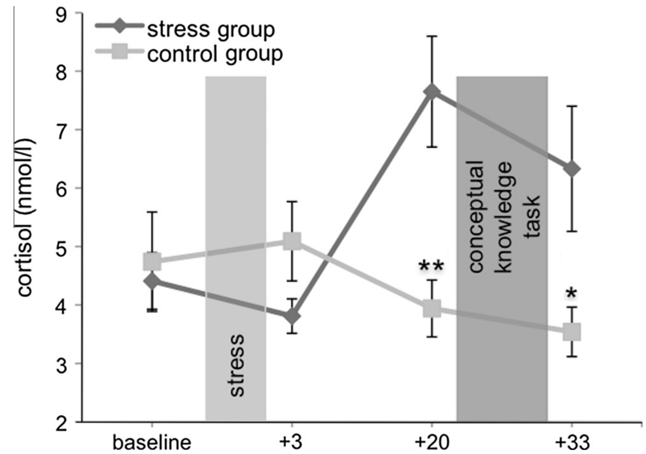

Fig. 1. Mean (±SEM) cortisol responses in the stress and the control group over the course of the experimental timeline (baseline, 3, 20 and 33 min after stress-onset). Experimental manipulation led to significantly increased cortisol concentrations in the stress compared to the control group, evident 20 (** $p < 0.01$) and 33 min (* $p < 0.05$) after stress-onset. The conceptual knowledge task took place between these two time points.

Analyses of blood pressure measures also showed that stress induction was successful (*systolic blood pressure*: time × group interaction, $F_{(1.8,68.6)} = 37.82$, $p < 0.001$, $\eta^2 = 0.50$, main effect of group, $F_{(1,38)} = 19.96$, $p < 0.001$, $\eta^2 = 0.34$, main effect of time, $F_{(1.8,68.6)} = 66.36$, $p < 0.001$, $\eta^2 = 0.64$; *diastolic blood pressure*: time × group interaction, $F_{(1.9,73.6)} = 59.23$, $p < 0.001$, $\eta^2 = 0.61$, main effect of group, $F_{(1,38)} = 11.18$, $p = 0.002$, $\eta^2 = 0.23$, main effect of time, $F_{(1.9,73.6)} = 76.97$, $p < 0.001$, $\eta^2 = 0.67$). Follow-up analysis revealed significantly higher blood pressure during (*systolic blood pressure*: $t_{(38)} = 7.00$, $p < 0.001$, $\eta^2 = 0.56$; *diastolic blood pressure*: $t_{(38)} = 7.37$, $p < 0.001$, $\eta^2 = 0.59$) and post hand immersion (only for *systolic blood pressure*: $t_{(38)} = 2.05$, $p = 0.047$, $\eta^2 = 0.10$; *diastolic blood pressure*: $t_{(38)} = 1.48$, $p > 0.05$) in the stress compared to the control group, but not at baseline (*systolic blood pressure*: $t_{(38)} = 1.43$, $p > 0.05$; *diastolic blood pressure*: $t_{(38)} = 0.62$, $p > 0.05$; cf. Table 1).

Furthermore, the stress group rated the hand immersion into water as significantly more stressful ($t_{(20.6)} = 4.99$, $p < 0.001$, $\eta^2 = 0.55$), painful ($t_{(19.9)} = 7.28$, $p < 0.001$, $\eta^2 = 0.73$) and unpleasant ($t_{(21.0)} = 7.40$, $p < 0.001$, $\eta^2 = 0.72$; cf. Table 1) compared to the control group.

3.3. Stress effects on conceptual knowledge retrieval

The sample mean accuracy was 77.0% (*SD* = 7.9; *min* = 58.0; *max* = 91.0) and the sample median reaction time 2433 ms (*SD* = 686; *min* = 1521; *max* = 5204). Faster reaction times were found for compatible compared to incompatible statements (main effect of compatibility: $F_{(1,38)} = 70.10$, $p < 0.001$, $\eta^2 = 0.65$). Importantly, compatibility influenced reaction times as a function of scientific correctness and stress (compatibility × scientific correctness × group interaction: $F_{(1,38)} = 6.94$, $p = 0.012$, $\eta^2 = 0.15$), no other significant main effects or interactions were detected (all $F < 2.13$, all $p > 0.05$). Post-hoc tests revealed that this interaction was not driven by compatible statements (all $F < 1.15$, all $p > 0.05$), but by incompatible statements (scientific correctness × group interaction: $F_{(1,38)} = 5.56$, $p = 0.024$, $\eta^2 = 0.13$): the control group showed faster reaction times on false-incompatible (e.g. "coats produce heat") compared to true-incompatible statements (e.g. "pressure produces heat"; $F_{(1,19)} = 8.16$, $p = 0.010$, $\eta^2 = 0.30$), whereas no such significant difference occurred in the stress group ($F_{(1,19)} = 1.39$, $p > 0.05$; cf. Fig. 3). An effect of heightened

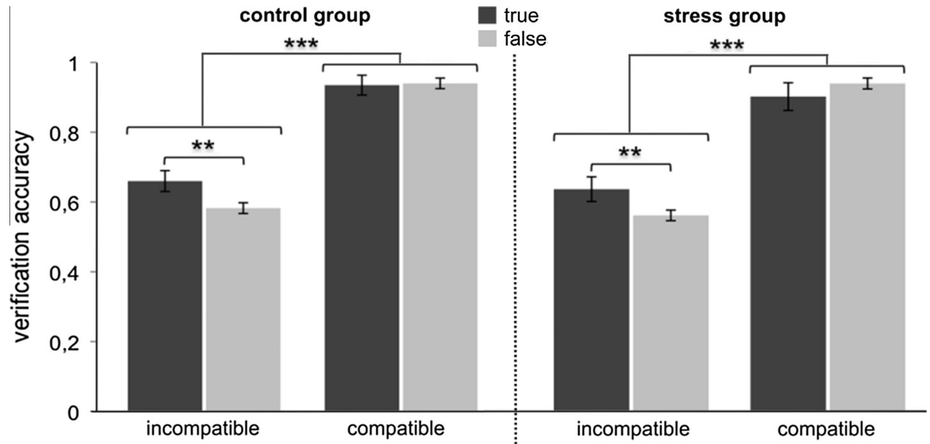

Fig. 2. Mean (±SEM) verification accuracy for statements as a function of compatibility (compatible vs. incompatible) and scientific correctness (true vs. false) in the stress and the control group. Across both groups, higher accuracy scores were obtained for compatible compared to incompatible statements ($^{***}p < 0.001$). Furthermore, scientific correctness modulated this compatibility effect as evidenced by higher accuracy scores for true-incompatible compared to false-incompatible statements ($^{**}p \leqslant 0.01$), but not for true- vs. false-compatible statements.

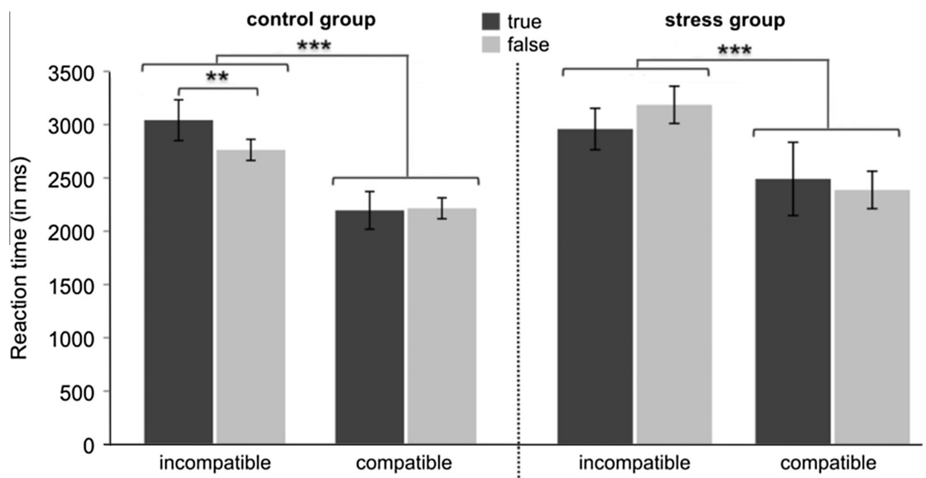

Fig. 3. Median (±SEM) reaction time (in ms) as a function of compatibility (compatible vs. incompatible) and scientific correctness (true vs. false) in the stress and the control group. Across both groups, faster reaction times occurred for compatible compared to incompatible statements ($^{***}p < 0.001$). In addition, stress modulated this compatibility effect in interaction with scientific correctness as evidenced by faster reaction times for false-incompatible compared to true-incompatible statements in the control group ($^{**}p \leqslant 0.01$), but not in the stress group.

stress hormones was also observed on the correlational level: significant positive associations occurred between cortisol increase and reaction times on false-incompatible (Δ cort: $r = 0.36$, $p = 0.024$; AUC$_i$: $r = 0.32$, $p = 0.046$) as well as false-compatible statements (Δ cort: $r = 0.37$, $p = 0.020$; AUC$_i$: $r = 0.36$, $p = 0.022$), even when controlling for the factor group (*false-incompatible statements*: Δ cort: $r = 0.33$, $p = 0.044$; *false-compatible statements*: Δ cort: $r = 0.36$, $p = 0.024$; AUC$_i$: $r = 0.35$, $p = 0.031$), but not for true statements (all $r < 0.28$, all $p > 0.05$). Thus, increases in cortisol secretion are associated with slower reaction times when evaluating scientifically false statements regardless of their compatibility with scientific concepts. The SNS stress response (Δ heart rate, Δ sys, Δ dia) and post hoc ratings of the experimental condition were not correlated with reaction times (all $r < 0.27$, all $p > 0.05$).

Higher accuracy scores were observed for compatible than for incompatible statements (main effect of compatibility: $F_{(1,38)} = 258.17$, $p < 0.001$, $\eta^2 = 0.87$), which was subject to a modulation by scientific correctness (compatibility × scientific correctness interaction: $F_{(1,38)} = 24.46$, $p < 0.001$, $\eta^2 = 0.39$; see Fig. 2). No other significant main effects or interactions were detected (all $F < 3.12$, all $p > 0.05$). Across the entire sample, participants were more accurate in verifying true-incompatible than false-incompatible statements ($F_{(1,38)} = 10.52$, $p = 0.002$, $\eta^2 = 0.22$), whereas there was no difference between false-compatible and true-compatible accuracy statements ($F_{(1,38)} = 3.54$, $p > 0.05$). The SNS and HPA axis stress responses as well as post hoc ratings of the experimental condition were not correlated with individual accuracy scores (all $r < 0.24$, all $p > 0.05$).

4. Discussion

In the present study, we investigated the influence acute psychosocial stress and the associated SNS and HPA axis activation have on conceptual knowledge retrieval. As expected, statements incompatible with naïve concepts generally provoked longer reaction times, reflecting a more effortful processing when incompatible concepts coexist in memory. Stress exerted a selective effect on efficiency when incompatible statements had to be verified in that it abolished reaction time differences between true and false statements without altering accuracy. Furthermore, individual differences were observed indicating that persons with a higher increase in cortisol concentrations also displayed slower reaction times when evaluating scientifically false statements, regardless of their compatibility with naïve theories learned early in life.

A part of our results replicates earlier findings by Shtulman and Valcarcel (2012). As in the original study, we found longer reaction times and a lower accuracy for incompatible compared to compatible statements about concepts. This conceptual replication of Shtulman's main finding is remarkable, as we only used a subset of Shtulman's items in order to reduce testing time and we translated Shtulman's English items into German for our data collection in Germany. The fact that we were able to replicate Shtulman's findings with translated items shows that the observed effects are reliable and might occur on the level of semantics, that is, the meaning of concepts, and are not due to language-specific item characteristics. Overall, these results add to a growing number of recent studies providing evidence for the coexistence of naïve and scientific concepts in individuals (Babai et al., 2010; Kelemen et al., 2013; Potvin et al., 2015; Schneider & Hardy, 2013). Naïve concepts coexisting and conflicting with scientific concepts (i.e. incompatible concepts) must be actively suppressed during retrieval of the correct knowledge. Since this inhibitory process represents an effortful cognitive task, it leads to longer reaction times for such statements.

In line with previous behavioral and brain imaging studies (e.g., Marques et al., 2009), we found support for the hypothesis that evaluations of correct and incorrect statements involve partly different cognitive processes. Unlike Shtulman and Valcarcel (2012), we did not find a main effect of statement correctness on reaction times. However, we did find a statistically significant two-way interaction between compatibility and correctness regarding accuracy and a three-way interaction between compatibility, correctness, and experimental condition regarding reaction time possibly concealing the main effect of statement correctness.

Acute psychosocial stress interacted with the retrieval of conceptual knowledge. In particular, when participants were evaluating incompatible statements, they were faster with scientifically false statements (e.g. "coats produce heat") than with scientifically correct statements (e.g. "pressure produces heat") in the control group but not in the stress group. Stress thus abolished the reaction time differences between true and false-incompatible items in our study, suggesting a comparably effortful reasoning process required to answer these statements. It is conceivable that the reasoning process after stress took more time because the retrieval of relevant conceptual knowledge was impaired by heightened cortisol concentrations. Indeed, our reported correlations are well in line with this view as they show positive associations of cortisol increases with false-compatible and -incompatible statements. The more circulating GCs are released, the more time it takes for a person to retrieve information from conceptual knowledge. This view corroborates previous studies on working memory performance under conditions of high cognitive load which observed longer reaction times after stress induction (Oei et al., 2006; Schoofs et al., 2008, 2009) or after GC-administration (Lupien, Gillin, & Hauger, 1999). Additionally, our results extend previous studies by showing that most likely high cortisol concentrations exerted a selective effect on the efficiency of the retrieval process during reasoning on false-incompatible statements (concerning reaction times), without, however, altering effectiveness in terms of accuracy performance.

On a conceptual level, our findings suggest that stress rather slowed down verifications for false-incompatible statements than speeded up verifications for true-incompatible statements. Thus, false-incompatible statements, which are in line with naïve concepts acquired early in life but conflict with scientific concepts acquired later, seem to be more prone to stress-related effects compared to true-incompatible statements. Possibly, naïve concepts still exist at a more habitual (or automatized and implicit) level which is usually inhibited by newly learned scientific concepts at a more conscious and controlled level. However, stress might have impaired this inhibition and could have triggered habitual processing as has been shown convincingly before (for a review: cf. Schwabe & Wolf, 2013). For true-incompatible statements, conscious and controlled processing seems to have overridden the habitual level, at least, it is stable enough to be insensitive to the detrimental effects of stress. Alternatively, stress might have slowed down verification for false-incompatible information by interfering with higher-order working processes (Schoofs et al., 2008, 2009), which are needed for false-incompatible more than for true-incompatible statements.

Previous studies have found negative effects of psychosocial stress on the retrieval of declarative memories of word lists (de Quervain et al., 2000; Kuhlmann et al., 2005) and biographical information (Merz et al., 2010). It might seem surprising that stress did not have a main effect on the retrieval of conceptual knowledge in the current study. However, this finding is explained by the fact that scientifically true-compatible statements (e.g.: "ovens produce heat.") and scientifically false-compatible statements (e.g., "rain produces heat.") are in line not only with scientific concepts but also with children's naïve concepts. If even children with little or no formal instruction on the topic can correctly evaluate the truth values of these statements, it can be assumed that the underlying concepts are highly overlearned and deeply rooted in everyday life for adults, including the participants in our study. The evaluation of compatible concepts might thus require such little cognitive resources that stress-induced impairments are minimal and hard to detect. Subsequent studies with high statistical power will be needed to evaluate whether the effects of stress on the retrieval of highly overlearned compatible concepts are entirely absent or just very small. Additionally, an exploration of interventions reducing the stress-induced effects needs further attention, such as behavioral (e.g. listening to music (Linnemann, Ditzen, Strahler, Doerr, & Nater, 2015) or timing of the stressor relative to retrieval (Schwabe & Wolf, 2014)) or pharmacolocial methods (e.g. administration of the beta-blocker propranolol (Schwabe et al., 2012)).

Limitations of the current study encompass the inclusion of male participants only. A consistent body of evidence indicates that sex differences influence the endocrine stress response (e.g. Kirschbaum, Kudielka, Gaab, Schommer, & Hellhammer, 1999; Kudielka et al., 2009), suggesting they might also impact conceptual knowledge retrieval performance under stress. Future studies should include women in different phases of the menstrual cycle and taking oral contraceptives (cf. Merz et al., 2012) to investigate how circulating sex hormone concentrations affect conceptual knowledge retrieval in interaction with stress. Furthermore, in a sample comprised of professional scientists, different response patterns were observed in terms of a higher accuracy and a reduced bias towards naïve assumptions compared to populations with a lower level of education (Kelemen et al., 2013). Since we studied a student population (with a broad range of subjects of study), no conclusions about moderating effects of the degree of domain-specific expertise on the interference of naïve and scientific concepts can be drawn, which calls for further exploration. It also awaits further clarification whether retrieval impairment is a general phenomenon, as stated by Shtulman and Valcarcel (2012), independent of domain even under stress or whether – under circumstances yet to be explored – domains are differentially affected depending on the degree of conceptual change. For example, it remains to be tested whether the effects are stable across domains differing in the amount, the interrelatedness, the difficulty, or abstractness of their content.

5. Conclusions

Taken together, the present results add further support to the finding of a detrimental effect of acute stress on memory retrieval. When participants evaluate statements inconsistent with naïve conceptions, retrieval efficiency is higher for scientifically true than for scientifically false statements under regular conditions, but not under the influence of stress. The impairment in efficiency is reflected by delayed reaction times, while effectiveness in the form of accuracy remains unaltered by stress. Positive associations between cortisol response and response times for scientifically false statements suggest that retrieval impairments are at least partly mediated by activation of the HPA axis. Naïve theories seem to coexist alongside scientific theories, as indicated by higher reaction times and fewer correct responses on statements about incompatible concepts. The fact that stress affects the cognitive processing of some components of conceptual knowledge but not others can help to better understand the retrieval of conceptual knowledge in real-life situations, including examinations and problem solving in the workplace. Future studies should aim to explore behavioral and pharmacological interventions to reduce these stress-induced effects.

Role of the funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

All authors declare no conflict of interest.

Acknowledgements

We gratefully acknowledge the help of Lisa Glänzer, Isabel Rausch, Lisa-Marie Stock and Shu Zhang during recruitment and stressing of participants. We thank the members of the biochemical laboratory of the Department of Biological and Clinical Psychology (University of Trier) for saliva analyses and Eve-Mariek Hessas for language editing. In addition, we acknowledge the helpful comments raised by the reviewers.

References

Alibali, M. W., & Sidney, P. G. (2015). Variability in the natural number bias: Who, when, how, and why. *Learning and Instruction, 37*, 56–61.
Babai, R., Sekal, R., & Stavy, R. (2010). Persistence of the intuitive conception of living things in adolescence. *Journal of Science Education and Technology, 19*, 20–26.
Brault Foisy, L.-M., Potvin, P., Riopel, M., & Masson, S. (2015). Is inhibition involved in overcoming a common physics misconception in mechanics? *Trends in Neuroscience and Education, 40*, 26–36.
Buss, C., Wolf, O. T., Witt, J., & Hellhammer, D. H. (2004). Autobiographic memory impairment following acute cortisol administration. *Psychoneuroendocrinology, 29*, 1093–1096.
Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.
Carey, S. (1992). The origin and evolution of everyday concepts. In R. Giere (Ed.), *Cognitive models of science* (pp. 89–128). Minneapolis: University of Minnesota Press.
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*, 313–314.
diSessa, A. A., Gillespie, N. M., & Esterly, J. B. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science, 28*, 843–900.
Dressendörfer, R., Kirschbaum, C., Rohde, W., Stahl, F., & Strasburger, C. J. (1992). Synthesis of a cortisol-biotin conjugate and evaluation as a tracer in an immunoassay for salivary cortisol measurement. *Journal of Steroid Biochemistry and Molecular Biology, 43*, 683–692.
Het, S., Ramlow, G., & Wolf, O. T. (2005). A meta-analytic review of the effects of acute cortisol administration on human memory. *Psychoneuroendocrinology, 30*, 771–784.
Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General, 142*, 1074–1083.
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*, 154–162.
Kudielka, B. M., Hellhammer, D. H., & Wüst, S. (2009). Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology, 34*, 2–18.
Kuhlmann, S., Piel, M., & Wolf, O. T. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *Journal of Neuroscience, 25*, 2977–2982.
Linnemann, A., Ditzen, B., Strahler, J., Doerr, J. M., & Nater, U. M. (2015). Music listening as a means of stress reduction in daily life. *Psychoneuroendocrinology, 60*, 82–90.
Lupien, S. J., Gillin, C. J., & Hauger, R. L. (1999). Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: A dose-response study in humans. *Behavioral Neuroscience, 113*, 420–430.
Machery, E. (2010). Précis of doing without concepts. *Behavioral and Brain Sciences, 33*, 195–244.
Marques, J. F., Canessa, N., & Cappa, S. (2009). Neural differences in the processing of true and false sentences: Insights into the nature of 'truth' in language comprehension. *Cortex, 45*, 759–768.
Merz, C. J., Tabbert, K., Schweckendiek, J., Klucken, T., Vaitl, D., Stark, R., & Wolf, O. T. (2012). Oral contraceptive usage alters the effects of cortisol on implicit fear learning. *Hormones and Behavior, 62*, 531–538.
Merz, C. J., Wolf, O. T., & Hennig, J. (2010). Stress impairs retrieval of socially relevant information. *Behavioral Neuroscience, 124*, 288–293.
Merz, C. J., & Wolf, O. T. (2015). Stress and emotional learning in humans: Evidence for sex differences. In R. M. Shansky (Ed.), *Sex differences in the central nervous system* (pp. 149–170). London, San Diego, Waltham, Oxford: Academic Press.
Neubauer, A. C., & Freudenthaler, H. H. (1994). Reaction times in a sentence-picture verification test and intelligence: Individual strategies and effects of extended practice. *Intelligence, 19*, 193–218.
Oei, N. Y. L., Elzinga, B. M., Wolf, O. T., de Ruiter, M. B., Damoiseaux, J. S., Kuijer, J. P., ... Rombouts, S. A. R. B. (2007). Glucocorticoids decrease hippocampal and prefrontal activation during declarative memory retrieval in young men. *Brain Imaging and Behavior, 1*, 31–41.
Oei, N. Y. L., Everaerd, W. T., Elzinga, B. M., van Well, S., & Bermond, B. (2006). Psychosocial stress impairs working memory at high loads: An association with cortisol levels and memory retrieval. *Stress, 9*, 133–141.
Perlman, W. R., Webster, M. J., Herman, M. M., Kleinman, J. E., & Weickert, C. S. (2007). Age related differences in glucocorticoid receptor mRNA levels in the human brain. *Neurobiology of Aging, 28*, 447–458.
Potvin, P., Masson, S., Lafortune, S., & Cyr, G. (2015). Persistence of the intuitive conception that heavier objects sink more: A reaction time study with different levels of interference. *International Journal of Science and Mathematics Education, 13*, 21–43.
Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology, 28*, 916–931.
Schlosser, N., Wolf, O. T., Fernando, S. C., Riedesel, K., Otte, C., Muhtz, C., ... Wingenfeld, K. (2010). Effects of acute cortisol administration on autobiographical memory in patients with major depression and healthy controls. *Psychoneuroendocrinology, 35*, 316–320.
Schneider, M., & Hardy, I. (2013). Profiles of inconsistent knowledge in children's pathways of conceptual change. *Developmental Psychology, 49*, 1639–1649.
Schnotz, W., & Preuß, A. (1997). Task-dependent construction of mental models as a basis for conceptual change. *European Journal of Psychology of Education, 12*, 185–211.
Schoofs, D., Preuß, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology, 33*, 643–653.
Schoofs, D., Wolf, O. T., & Smeets, T. (2009). Cold pressor stress impairs performance on working memory tasks requiring executive functions in healthy young men. *Behavioral Neuroscience, 123*, 1066–1075.
Schwabe, L., Haddad, L., & Schächinger, H. (2008). HPA axis activation by a socially evaluated cold-pressor test. *Psychoneuroendocrinology, 33*, 890–895.
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.
Schwabe, L., & Wolf, O. T. (2013). Stress and multiple memory systems: From 'thinking' to 'doing'. *Trends in Cognitive Sciences, 17*, 60–68.
Schwabe, L., & Wolf, O. T. (2014). Timing matters: Temporal dynamics of stress effects on memory retrieval. *Cognitive, Affective, & Behavioral Neuroscience, 14*, 1041–1048.
Schwabe, L., Wolf, O. T., & Oitzl, M. S. (2010). Memory formation under stress: Quantity and quality. *Neuroscience and Biobehavioral Reviews, 34*, 584–591.
Shtulman, A., & Harrington, K. (2016). Tensions between science and intuition across the lifespan. *Topics in Cognitive Science, 8*, 118–137.

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition, 124*, 209–215.
Takahashi, T., Ikeda, K., Ishikawa, M., Tsukasaki, T., Nakama, D., Tanida, S., & Kameda, T. (2004). Social stress-induced cortisol elevation acutely impairs social memory in humans. *Neuroscience Letters, 363*, 125–130.
Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*, 45–69.
Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*, 535–585.
Wolf, O. T. (2009). Stress and memory in humans: Twelve years of progress. *Brain Research, 1293*, 142–154.
Yee, E., Chrysikou, E. G., & Thompson-Schill, S. L. (2013). Semantic memory. In K. N. Ochsner & S. Kosslyn (Eds.). *The Oxford handbook of cognitive neuroscience* (Vol. 1, pp. 353–374). Oxford: Oxford University Press.