Contemporary Educational Psychology 49 (2017) 238–249

Contents lists available at ScienceDirect

Contemporary Educational Psychology

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

Empirical study

Stress and emotions during experiments in biology classes: Does the work setting matter?

Nina Minkley [a,*,1], Tobias Ringeisen [b,1], Lukas B. Josek [a], Tobias Kärner [c]

[a] Ruhr-Universität Bochum, Faculty of Biology and Biotechnology, Behavioural Biology and Biology Education, Universitätsstr. 150, 44780 Bochum, Germany
[b] Merseburg University of Applied Sciences, School of Business, Eberhard-Leibnitz-Str. 2, 06217 Merseburg, Germany
[c] University of Bamberg, Chair of Business Education and Educational Management, Kärntenstraße 7, 96052 Bamberg, Germany

ARTICLE INFO

Article history:
Available online 10 March 2017

Keywords:
Subjective stress
Emotions
Physiological stress
Experiments

ABSTRACT

Experiments are a complex teaching method carrying a high cognitive load and the risk of failure, which both may induce stress among students. However, it remains unclear if the work setting modulates physiological, subjective, and/or emotional stress responses during experiments. In a randomized experimental field study school students ($N = 104$) either watched a biology experiment on video (passive condition), conducted the experiment on their own (active condition) or in small groups (interactive condition). Meanwhile, their subjective stress perception, heart rate variability (HRV), salivary cortisol concentration, and achievement emotions were assessed. In the active condition we observed the strongest subjective and HRV stress responses, followed by the interactive condition. Students of the passive condition displayed the weakest stress reactions. Students of the other two conditions showed a weakened diurnal cortisol decrease, indicating more stress. Across conditions, enjoyment dropped and boredom increased, most pronounced in the passive condition. Moreover, there were some associations between subjective, emotional and physiological stress responses. The findings suggest that conducting experiments alone carries the risk of self-attributed failure signified by elevated stress. In contrast, conducting an experiment in a group is less stressful, as others may constitute a source of support. Watching others conduct an experiment carries a low risk of failure and, thus, the lowest stress responses, but comes with the cost of minimized enjoyment and maximized boredom.

1. Introduction

Experiments are an important learning method in science education, as they enable students to acquire scientific knowledge, develop scientific reasoning skills (Engelmann & Fischer, 2014), obtain insights into scientific methods, and develop problem-solving competencies (Barzel, Reinhoffer, & Schrenk, 2012; for a review, see e.g., Hofstein & Lunetta, 2003). Despite their benefits, many students express dislike about experiments up-front, especially during adolescence, but tend to become interested once they have mastered hands-on experiments themselves (Carmona Miranda, 2012; Dohn, Fago, Overgaard, Madsen, & Malte, 2016). A reason for the disapproval could be the fact that experimental work settings constitute complex learning environments that carry the risk of failure and negative performance evaluations and may, thus, elicit bodily-affective responses, in particular the experience of stress, discrete emotional states and accompanying physiological responses (Alsop & Watts, 2003; Engelmann & Fischer, 2014; Hofstein & Lunetta, 2003; Pekrun & Stephens, 2012). The consequences may be costly, as aversive emotions may predict lower engagement and performance in science subjects (Randler, Hummel, & Wüst-Ackermann, 2013; Wigfield, Tonks, & Klauda, 2009), and may prevent students from pursuing studies or a career in science (Chow, Eccles, & Salmela-Aro, 2012; Osborne & Dillon, 2008).

Research within the context of regular classroom settings suggests that the experience of discrete emotional states, and thus

Abbreviations: AEQ, Achievement Emotions Questionnaire; ANOVA, Analysis of Variance; AUCg, area under the curve with respect to ground; AUCi, area under the curve with respect to increase; ANS, autonomic nervous system; BMI, body mass index; CVT, Control-Value Theory; DNA, deoxyribonucleic acid; HRV, heart rate variability; HPA-axis, hypothalamus pituitary adrenal axis; LF/HF, low frequency/high frequency; OC, oral contraceptives; rmsSD, root mean square of successive differences; VAS, Visual Analogue Scale.

* Corresponding author.
E-mail addresses: nina.minkley@rub.de (N. Minkley), tobias.ringeisen@hs-merseburg.de (T. Ringeisen), lukas.josek@ruhr-uni-bochum.de (L.B. Josek), tobias.kaerner@uni-bamberg.de (T. Kärner).

[1] N. Minkley and T. Ringeisen shared first authorship.

http://dx.doi.org/10.1016/j.cedpsych.2017.03.002

their effects on performance, depend on a number of situational characteristics (e.g., Pekrun, 2006; Pekrun & Stephens, 2012). For the context of science education, there is evidence that the features of the learning environment (e.g., control over learning activities) and the interaction with teachers and peers (e.g., participation in class) trigger different emotions, which in turn affect the use of learning strategies, achievement motivation, class engagement, and academic performance (e.g., Bellocchi, Quigley, & Otrel-Cass, 2017; Sinatra, Broughton, & Lombardi, 2014). Only few studies have investigated emotions when students perform experiments themselves, mostly focusing on negative emotions like anxiety in physics (Gungor, Eryilmaz, & Fakioglu, 2007) or disgust in biology (Randler et al., 2013). A recent study examined both negative and positive emotions: When students conducted experiments, they experienced high enjoyment and low boredom, and reported increasing situational competence in handling the topic (Itzek-Greulich et al., 2017). However, it remains largely unknown which characteristics of the work setting, especially if independent from a specific content, are crucial in shaping the students' bodily-affective responses (cf. Alsop & Watts, 2003; Pekrun, 2006). First, the available studies mostly addressed negative emotions, neglecting the role of positive affect. Second, existing research on situational antecedents of emotions has seldom relied on a taxonomy of work settings which specifies the constellation of relevant situational features. Third, previous studies have rarely used a randomized experimental design.

Using Control-Value Theory of achievement emotions in learning settings (CVT; Pekrun, 2006) and Chi' taxonomy of work settings (Chi, 2009; Menekse, Stump, Krause, & Chi, 2013), the current study investigated perceived stress, positive as well as negative emotions, and accompanying physiological correlates in high school students who conducted a biology experiment under three constructive conditions, which were enriched either with passive, active, or interactive elements. All conditions included the same inquiry-based elements requiring the students to formulate hypotheses, record the results, and reflect upon them. The experiment concentrated on the domain of biology for two reasons: Many high school graduates suffer from fragmented knowledge, and thus struggle in introductory biology courses at college (Harackiewicz et al., 2014). On the other hand, domain-specific knowledge in biology can easily be built up, as previous research identified a link between affective factors and learning especially when students are confronted with inquiry-based approaches like experiments (Dohn et al., 2016). The findings of our study may help to design hands-on experiments that induce low to moderate levels of (activating) stress and positive emotions like enjoyment, but reduce negative emotions like boredom which may have positive effects on students' performance and their attitudes towards science.

1.1. Perceived stress and emotions in academic settings

In learning settings, students may encounter different levels of perceived stress and bodily-affective responses (Alsop & Watts, 2003; Pekrun, 2006; Pekrun & Stephens, 2012). Appraisal emotion theories define stress as a temporary psychological state accompanied by intense arousal which emerges if a situation or a task is appraised as challenging or threatening (Lazarus, 1991). When individuals consider a stressful situation to exceed their coping capabilities, threat appraisals occur. In response, debilitating stress and aversive negative emotions like anxiety may arise. If the person appraises the stressor as challenging the demands are perceived as manageable. Hence, activating levels of stress are evoked which are often accompanied by positive emotions like enjoyment.

Introducing CVT, Pekrun and colleagues (Pekrun, 2006; Pekrun & Stephens, 2012) refined these propositions for the context of learning settings. According to CVT, students appraise potentially stress-inducing characteristics of the educational environment on two appraisal dimensions: (1) subjective control over achievement activities and outcomes and (2) the value attached to them. Achievement activities refer to all behaviors that enable students to deliver academic performance. Examples comprise dealing with external demands and performance evaluations (e.g., when giving a presentation), self-regulation (e.g., during learning), or social interactions with peers (e.g., in study groups) and teachers (e.g., during class). The experience of control is determined by prospective mastery expectations like self-efficacy beliefs as well as retrospective attributions. Value appraisals, on the other hand, signify whether students rate these activities and outcomes as positive or negative and the extent to which they are considered as personally relevant.

Depending on the situational characteristics, and subsequent control and value appraisals, discrete positive (e.g., hope or enjoyment) and/or negative emotions (e.g., anxiety or boredom) may co-occur. Emotions are conceptualized as a coordinated set of affective, cognitive, physiological, and behavioral components (Pekrun & Stephens, 2012). Achievement emotions are defined as emotions that arise in relation to achievement activities and/or subsequent outcomes. When students conduct experiments, predominantly activity-related achievement emotions are expected to occur, with enjoyment (positive, high activation) and boredom (negative, low activation) being the most significant. On the other hand, anxiety represents the most important outcome-related achievement emotion (Itzek-Greulich et al., 2017; Muis et al., 2015; Pekrun & Stephens, 2012).

1.2. Physiological stress correlates in academic settings

Besides triggering emotions and subjective stress, academic settings may also elicit physiological stress responses that can be measured through increases in the cortisol concentration and changes in heart rate variability (HRV) measurements (e.g., Hjortskov et al., 2004; Minkley & Kirchner, 2012). When someone faces a stressor, the secretion of cortisol is activated via the hypothalamus pituitary adrenal axis (HPA-axis), leading to a cortisol increase in the saliva, which peaks about 20 min later. Also, an attenuation of the natural diurnal cortisol decrease, which is characterized by a relatively stable peak concentration about 30 min after awakening and a continuous decrease thereafter, can be interpreted as a stress response (Foley & Kirschbaum, 2010; Hellhammer, Wüst, & Kudielka, 2009; Kirschbaum, Tietze, Skoluda, & Dettenborn, 2009). The majority of research has focused on acute stress during examinations, which were identified as social-evaluative stressors that lead to an increase in the cortisol concentration, at least directly before the examination (e.g., Dickerson & Kemeny, 2004; Spangler, Pekrun, Kramer, & Hofmann, 2002). Although, such an acute hormonal stress response can lead to allostasis and adaption (McEwen, 1998), chronic elevated cortisol concentrations can cause serious health problems (McEwen, 1998; Noll & Kirschbaum, 2006; Rensing, Koch, Rippe, & Rippe, 2006) and impair memory retrieval (Wolf, 2009), which may influence academic performance.

Besides, the autonomic nervous system (ANS) is activated if someone is confronted with a stressor. In contrast to the HPA-axis, the ANS ensures a much more rapid adaptation to stressful situations as it is based on nervous signaling instead of a hormonal cascade (Baert, Casier, & De Raedt, 2012; Kemeny, 2003). The ANS can be subdivided into two subsystems: Under stress, the sympathetic nervous system enables the organism to adapt rapidly to demanding situations, while the parasympathetic nervous system controls physiological resting functions (Kemeny, 2003). The antagonistic relation between these systems can be measured by changes in heart rate variability. The ratio between low and high frequencies (LF/HF ratio) is a common *frequency domain measure* of HRV mirroring sympathovagal balance (Sleight & Bernardi, 1998). In that regard, the high frequency domain (HF: 0.15–0.4 Hz) primarily reflects the activity of the parasympathetic nervous system and the low frequency domain (LF: 0.04–0.15 Hz) the activity of the sympathetic nervous system (Hjortskov et al., 2004; Malik et al., 1996). Besides the LF/HF ratio, the root mean square of successive differences (rmsSD, [ms]) represents a common *time domain measure* of HRV and is associated with parasympathetic activity (Malik et al., 1996). Under stress, the LF/HF ratio increases (e.g., Isowa, Ohira, & Murashima, 2006), while the rmsSD decreases (e.g., Malik et al., 1996).

1.3. Associations between stress, emotions, physiological stress responses, and performance

Some studies have examined associations between subjective and physiological stress correlates, emotions, and performance, in naturalistic academic settings. Empirical evidence is mixed and limited to examinations (for a review, see Campbell & Ehlert, 2012; Zeidner, 2007). While some authors reported significant associations between cortisol concentrations, emotions, and subjective stress (e.g., Lindahl, Theorell, & Lindblad, 2005; Spangler et al., 2002) which might go along with debilitated performance due to impairments in working memory capacity, retrieval and processing of information (Spering, Wagener, & Funke, 2005; Wolf, 2009), others could not find them (e.g., Minkley, Westerholt, & Kirchner, 2014). No study, to our knowledge, has investigated the associations between different physiological stress parameters with regard to educational phenomena in science education.

In order to disentangle the relationships between achievement emotions, perceived stress, physiological stress, and performance, assumptions and findings based on related cognitive emotion theories are integrated within the framework of CVT in the following. Specifically, we considered recent research reviews with regard to the Yerkes-Dodson-law (Hanoch & Vitouch, 2004), flow theory (Nakamura & Csikszentmihalyi, 2009), and cognitive load theory (Sweller, Ayres, & Kalyuga, 2011): In essence, an inverted U-shaped relation between subjective and physiological stress, and performance is expected which is emphasized by different patterns of control/value appraisals and discrete emotions. For instance, there is evidence for such a relationship between the cortisol concentration and recall performance: The best performances were recorded during a moderate elevation of salivary cortisol (Schilling et al., 2013). However, there is no "optimal" cortisol concentration because of strong interindividual differences with regard to the cortisol response (Dougan, Hastings, Granger, & Usher, 2001; Szalma, 2012).

In terms of achievement emotions, enjoyment is expected to co-occur with low to moderate levels of stress if the performance criteria are clear and students perceive the learning activities as challenging as well as both controllable and personally important, e.g. due to high competence beliefs. Up to an intermediate level, increasing stress is thus functionally activating as it enables students to engage in compensatory effort investment, which increases their ability to perform. Under such conditions, stress should be positively related to enjoyment, which in turn, serves as a positive predictor of learning outcomes such as high intrinsic achievement motivation, an effective use of learning strategies, greater class engagement, and better academic performance (Cassady & Johnson, 2002; Itzek-Greulich et al., 2017; Pekrun & Stephens, 2012).

Conversely, boredom is instigated when the activity lacks incentive value and/or does not constitute a challenge. Under such conditions, low levels of stress are linked to moderate to high levels of boredom and low performance. Intense boredom, however, is perceived as stressful and shows positive relations with stress (Alsop & Watts, 2003; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011; Pekrun & Stephens, 2012). Should the students' performance be evaluated (e.g., during or after an experiment by supervising teachers), intense feelings of stress and anxiety are expected to emerge, especially when students experience uncertainty and low control about the achievement outcomes, while they are considered personally significant (Pekrun, 2006; Pekrun & Stephens, 2012). Thus, if the conditions are characterized by high personal relevance, high uncertainty and a low sense of control, stress levels may rise above an intermediate level – particularly in terms of physiological arousal and worries about failure – which may undermine intrinsic achievement motivation, increases mental load, diminish engagement in academic tasks, and debilitate performance (Cassady & Johnson, 2002; Macher, Paechter, Papousek, & Ruggeri, 2012; Petermann & Winkel, 2007). In summary, U-shaped relations are proposed between stress intensity and emotions. Low to moderate stress levels should be positively related to enjoyment and negatively to boredom. Above an intermediate stress level, moderate to high stress response should be negatively related to enjoyment but positively to boredom (Pekrun & Stephens, 2012).

1.4. Stress and emotions when students conduct experiments in science education

A few studies have investigated stress experiences and emotions when students perform experiments in the context of science education, mostly focusing on negative emotions like boredom, anxiety, or disgust (Gungor et al., 2007; Itzek-Greulich et al., 2017; Randler et al., 2013). These studies examined the emotion-inducing effect of selected activities and/or situational characteristics. However, it remains largely unknown which combination of work characteristics regardless of the content are crucial in shaping the students' bodily-affective responses when they conduct experiments (Alsop & Watts, 2003; Pekrun, 2006; Pekrun & Stephens, 2012). Keeping in mind that the learning content in science education in German schools is mainly determined by the curriculum, and thus by the Ministry of Education, teachers have limited degrees of freedom in their scope of action to influence students affective responses with regard to a concrete learning topic (Osborne & Dillon, 2008). Teachers have to handle a defined topic, but can decide on the didactical design they use in education. Therefore, we kept the learning content constant and focused on the effects of the didactical work setting on students' stress responses. We applied Chi's taxonomy of work settings (Chi, 2009; Menekse et al., 2013) which integrates different work characteristics like subjective control, active engagement and peer interaction that other studies identified as predictors of emotions in science education (Bellocchi et al., 2017; Sinatra et al., 2014).

Experiments can be classified with regard to their level of inquiry. In essence, they range from low inquiry settings with strict instructions to high inquiry open-ended experiments, with several intermediate stages (for an overview, see Priemer, 2011). In line with Chi and colleagues (Chi, 2009; Menekse et al., 2013), passive, active, constructive, and interactive work settings (and combinations thereof) may be implemented when students conduct an experiment. In passive settings, students simply observe others conducting an experiment. In active settings, students conduct the experiment themselves, which requires them to physically manipulate the task, retrieve existing knowledge, and encode new information. In constructive settings, in addition, students have to reflect upon and interpret results and integrate new information with existing knowledge. In interactive settings, students conduct experiments in dyads or small groups, which involve intensive interaction. Interactive settings are assumed to be the most effective in terms of acquiring scientific knowledge and reasoning, followed by constructive settings, active settings, and passive settings (Chi, 2009; Menekse et al., 2013). When students conduct experiments at school, usually a constructive format is realized. However, passive, active, or interactive features may be added to accentuate the quality of the setting (Hofstein & Lunetta, 2003; Priemer, 2011).

Comparing the work settings as introduced by Chi (2009), the following patterns in perceived stress and emotions are to be expected (Pekrun, 2006; Pekrun, Goetz, Daniels, Stupnisky, & Perry, 2010). In passive settings, students face a low risk of failure and, thus, hardly a threat to their self-worth. Levels of subjective and physiological stress as well as anxiety should be low. Due to a lack of practical involvement, boredom should be high and enjoyment low (Pekrun et al., 2010). In active and constructive settings, students usually work individually. The advantage herein is that all students can work at their own speed, and with varying approaches (Berck & Graf, 2010; Rinschede, 2007). Under such circumstances, enjoyment should be high and boredom low (Pekrun et al., 2010). However, as individual performances and potential failure are salient to teachers, perceived and physiological stress responses are expected to be moderate (Pekrun, 2006). In interactive peer settings, students work together, which may enhance productivity and mutual intellectual stimulation, reduce individual mental effort, as well as balance out knowledge gaps and working memory load (Gillies, 2004; Sweller et al., 2011; Zhang, Ayres, & Chan, 2011). On the cost side, collaborative efforts require time and coordination, which may lead to lowered effectiveness and/or the disengagement of individual students (Engelmann & Fischer, 2014; LoGiudice, Pachai, & Kim, 2015). Thus, enjoyment is expected to be higher than in passive, but lower than in active or constructive work settings. The opposite pattern is expected for boredom (Pekrun, 2006; Pekrun et al., 2010). Conversely, perceived and physiological stress levels in interactive peer settings should be lower than those in active or constructive settings, but higher than those in passive settings (Pekrun & Stephens, 2012).

1.5. Current study

The current study compared subjective and physiological stress responses, as well as achievement emotions, when students worked on an experiment under three different constructive conditions: All students were required to formulate hypotheses, record the results, and reflect upon them, while conducting the experiment either individually (active condition), in small groups (interactive condition), or by watching peers conducting the experiment (passive condition). In the current study, neither value nor control appraisals were manipulated. However, in line with the assumptions of CVT on the interplay of situational features, appraisals, and stress parameters (Campbell & Ehlert, 2012; Pekrun & Stephens, 2012), we considered the three conditions as low to moderately stressful and emotion-inducing situations. Based on the above-mentioned research findings, the following hypotheses were evaluated:

Hypothesis 1: experiments cause different stress responses, depending on the work setting. We expected that conducting an experiment constitutes a low to moderate stressor for students, which is indicated by a change in physiological parameters (a rise in cortisol or at least an attenuated diurnal decrease; changes in HRV, as indicated by an increased LF/HF ratio and a decreased rmsSD) and emotional parameters (increasing subjective stress, decreasing enjoyment, and increasing boredom) during the course of the experiment.

Furthermore, we hypothesized that the experimental work setting modulates the students' stress responses. Based on Chi's (2009) taxonomy of work settings, we expected the stress responses to be most pronounced in the active condition, followed by the interactive condition, and the passive condition. In terms of activity achievement emotions, we hypothesized that enjoyment is higher in the interactive condition than it is in the passive condition, but lower than it is in the active condition. The opposite pattern was expected for boredom.

Hypothesis 2: different stress parameters are associated with each other. With regard to the different stress responses, we hypothesized subjective stress levels to be positively associated with cortisol levels and the LF/HF ratio, and negatively with the rmsSD. Moreover, we expected U-shaped relations between the intensity of stress responses and emotions. Since we considered that conducting an experiment constitutes a low to moderate stressor instigating low to moderate stress levels, we hypothesized boredom to be negatively associated with (1) the LF/HF ratio, (2) the cortisol concentration, and (3) perceived stress, but (4) positively with rmsSD values. Conversely, we expected a positive relationship between low to moderate stress levels and enjoyment.

2. Material and methods

2.1. Design and procedures

The present study is a randomized field experiment with two experimental groups and one control, conducted in the teaching and learning laboratory at the Ruhr-Universität in Bochum, Germany. The attending students participated in a one-day project on molecular biology, which required them to perform hands-on experiments. As the participants were recruited from six different schools, they attended the project on six different days.

Upon arrival at about 9:15 AM, the participants filled in questionnaires requesting demographic (sex, age), medical (weight, height, average wake-up time during the last week, chronic diseases, medication intake) and academic information (frequency of experiments in biology lessons, biology grade, previous knowledge about the cell structure and the construction of an experiment). We controlled the possibly confounding influence of these variables on cortisol levels and the interplay of the study variables in the data analyses.

One hour before the experiment started, participants were instructed to refrain from eating or drinking, in order to eliminate possible contamination of saliva cortisol measures (Foley & Kirschbaum, 2010). The actual experiment started at about 2:00 PM, after the students had participated in the molecular biology project. Then, the experimental task, the procedure and the objectives were explained to them. Based on Chi's (2009) taxonomy of work settings, students of each of the six schools were randomly assigned to one of the following constructive conditions: (1) an active condition (students conducted the experiment by themselves), an interactive condition (students conducted the experiment in groups of four), and a passive condition (students watched a film showing the same experiment performed by unfamiliar students).

During the experimental work that lasted 35 min at maximum, the students isolated DNA out of kiwi fruits. An adapted version of a well-established protocol (Scharfenberg, 2010) was used for different reasons: The experiment can easily be implemented in a laboratory with students, it avoids contact with dangerous materials, and it does not take long. Furthermore, the protocol increases student's sense of personal control, as it integrates well with regular biology high school curricula, and deepens students' understanding of already familiar topics. Moreover, the experimental environment was structured which reduces the risk of confounding factors to occur (e.g. false mixture of reagents). The experiment combined instructive and inquiry-based elements. This form of a partial inquiry-based experiment was chosen to ensure an age-adequate level of complexity and to allow the students an intensive processing of the experiment. Open-ended experiments, on the other hand, seemed inappropriate as they require high levels of content knowledge and experimental competencies which students of the respective age group often lack (Arnold, Kremer, & Mayer, 2013, 2014).

Before students started working in the three conditions, they were instructed how to practically implement the experiment. Also, they were provided with an instruction manual including the required materials and procedures. The students applied four types of everyday reagents: (a) dishwashing liquid, (b) salt, (c) dishwashing liquid and salt in combination, and (d) pure water as control. Subsequently, students crushed the cell membranes to extract the DNA out of the nucleus. After filtrating, the DNA could be precipitated, and thus, became optically visible. Students were instructed to investigate, which one of the four different solvents worked best to destroy the cell membrane. In terms of inquiry-based elements, the students in all three conditions had to hypothesize about the results upfront, test all of the solvents, observe the precipitation of DNA, and record their observations. After the experiment, students were asked to interpret the results with regard to their hypothesis. Thereafter, the correct result was shown and discussed with all participating students. The investigator observed the students continuously during their experiment, but did not interact with them.

Saliva samples for determining cortisol concentration and subjective stress ratings were recorded before, in the middle of, and after the experiment. Furthermore, the HRV of the students was measured continuously. Activity achievement emotions as well as knowledge about the cell membrane and the construction of an experiment were assessed by means of a paper-and-pencil questionnaire before and after the experiment. In particular, the students had to constitute or evaluate an experiment for the investigation of a given biological phenomenon, comparable to the tasks designed by Arnold et al. (2013).

2.2. Participants

The participants ($N = 104$; 49 boys, 55 girls) were high school students with a mean age of 17.48 years (±0.98 *SD*), who were recruited from six secondary schools in the Ruhr area in Germany. Since the experiment comprised complex inquiry-based elements requiring knowledge of DNA and cell structures, we chose high school students who had already covered the topic of the experiment in their biology curriculum. The distribution of male and female participants across the three conditions was balanced. There were no significant differences regarding age, body mass index (BMI), pre-experiment microbiology knowledge, and frequency of experiments in biology between the treatment groups (all p's > 0.50; univariate ANOVAs). Because of possible confounding effects on cortisol concentrations, the saliva of students with the presence of a serious medical condition, the use of any long-term medication (except oral contraceptives [OC]), or over-/underweight (BMI > 25 kg/m^2 or <18.5 kg/m^2) was not analyzed (Foley & Kirschbaum, 2010). The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Commission of the local medical school. All participants provided written informed consent.

2.3. Measurements

2.3.1. Heart rate variability (HRV)

In order to measure the cardiovascular activity of students we made use of HRV as a well-established quantitative marker of ANS activity (Agelink et al., 2001; Malik et al., 1996). We operationalized HRV by means of two measures. First, we used the LF/HF ratio which represents a frequency domain measure that is based on high frequency (HF: 0.15–0.4 Hz) components reflecting the activity of the parasympathetic nervous system, and low frequency (LF: 0.04–0.15 Hz) components reflecting the activity of the sympathetic nervous system. Second, we used the rmsSD which represents a time domain method for evaluating HRV (Hjortskov et al., 2004; Malik et al., 1996). The cardiovascular activity of the students was measured continuously during the experiment via a chest belt and storage devices that wirelessly recorded the data on an integrated memory chip (Medeia Ltd.®, cf. Danev, 2010; Kärner, 2015). The LF/HF ratio, as well as the rmsSD, was calculated through the Qhrv Assessment (Medeia Ltd.®) software.

2.3.2. Cortisol assessment

Participants' saliva was collected via a shortened straw in polypropylene micro tubes (SafeSeal, Sarstedt). Thereafter, the samples were frozen at −20 °C, thawed, vortexed (MS1, IKA), and centrifuged for 15 min at 2500*g* (Function Line 400R, Heraeus) twice. On the day of the analysis, the supernatant was transferred in duplicate into a pre-coated microwell plate and salivary cortisol was quantified using an immunoassay kit (IBL, Hamburg, Germany). Since blood contamination affects subsequent measurements, blood-contaminated samples were discarded (Westermann, Demir, & Herbst, 2004). Analyses were conducted using a 96-well ELISA reader (Thermo Fisher). Intra-assay coefficients of variance were below 6%, and inter-assay coefficients were below 11%.

2.3.3. Subjective stress perception

A Visual Analogue Scale (VAS, Luria, 1975) was used to assess subjective perceptions of stress. The VAS is a 100 mm-long horizontal line with a left end that is labeled, "no stress," and the right end, "maximum stress." Participants placed a cross on this scale, to express how stressed they felt at each point of measurement. Thereafter, the level of subjective stress was determined by measuring the distance between the left (no stress) end of the scale and the cross.

2.3.4. Emotions

Based on previous research showing that achievement emotions strongly differ by subject (e.g., Math or German) and setting-related activities (e.g., participation in class or engaging in group work) (Goetz et al., 2012), we used a domain- and activity-specific version of the Achievement Emotions Questionnaire (AEQ, Pekrun et al., 2011). Since the wording of the original measure refers to class-related emotions, we modified the item wording slightly to assess emotions regarding conducting experiments in biology, as suggested by the authors (Goetz et al., 2012). Parallel item wordings were used to assess boredom and enjoyment, with four items per scale. To account for the varying time frame of reference, emotions before the exam were worded prospectively (e.g., "I am going to enjoy conducting a biology experiment," for enjoyment) and retrospectively after the experiment (e.g., "I enjoyed conducting a biology experiment;" see Goetz et al., 2012, for detailed information). Responses to all items were evaluated on a Likert scale ranging from 1 ("*strongly disagree*") to 4 ("*strongly agree*"). Cronbach's alphas were 0.87 for boredom on both measurement occasions, and 0.88 (t1) and 0.80 (t2) for enjoyment.

2.3.5. Knowledge test

In order to determine the students' knowledge before and after conducting the experiment, the students had to solve seven tasks with regard to the construction of scientific experiments, the DNA structure, and DNA extraction methods. If a task was solved correctly the students received one point. If a task was partly solved the students received 0.5 points. Thus, the students could receive seven points in total.

2.4. Statistical analysis

All statistical analyses were performed using SPSS Version 22. Before the data were analyzed, a number of cases had to be excluded. With regard to cortisol, the saliva of four participants was contaminated with blood. In addition, eight participants had to be excluded, due to consumption of chronic medication, overweight, and underweight, respectively. A few HRV parameters of three individuals were missing due to measurement errors. Consequently, we examined saliva samples from 92 and HRV from 101 school students.

As the HRV was measured continuously over the course of the experiment, we aggregated the LF/HF ratio and the rmsSD to five-minute intervals and analyzed changes using a repeated-measures Analysis of Variance (ANOVA), with time as a repeated within-subject factor (seven aggregated points of measurement) and the work setting as the independent variable. For participants' cortisol concentrations and the subjective perception of stress, we calculated the area under the curve with respect to increase (AUCi). Here, the area below the cortisol concentration curve is calculated with respect to the baseline concentration, in order to achieve a summary measure of stress from samples collected over time (Pruessner, Kirschbaum, Meinlschmid, & Hellhammer, 2003).

Thereafter, ANOVAs with the work setting as the independent variable (three conditions: active, interactive, and passive) were conducted. Moreover, changes in achievement emotions after the experiment, as compared to before, were analyzed by means of repeated-measures ANOVAs.

To enable a comparison of association patterns among stress parameters across the three conditions, we calculated the area under the curve representing the course of the various stress parameters with respect to ground (area under the curve with respect to ground, AUCg; Pruessner et al., 2003) as a measure of total intensity versus time-dependent change for all physiological parameters, subjective stress, and the achievement emotions.

In the second step, correlational analyses were performed between the AUCg scores for the full sample and for each work condition separately. The associations between possibly confounding background variables (see section "design and procedures"; Hellhammer et al., 2009; Jennings & Gianaros, 2007; Pekrun & Stephens, 2012) and the study variables were analyzed separately by means of bivariate correlations. Their influence was then controlled for when computing the partial correlations among the study variables. When the sphericity assumption was violated, Greenhouse-Geisser-adjusted p-values were reported. The analyses include the partial η^2 as a measure of effect size.

3. Results

3.1. Frequency of experiments in biology education and changes in knowledge

In total, 31% of the students reported to have conducted experiments sometimes during biology lessons, while almost two thirds (65%) reported to have seldom or never conducted experiments. Only 3% had conducted experiments often. Before conducting the experiment, the students in the three conditions did not differ with regard to previous knowledge ($F(2, 89) = 0.44$, $p = 0.65$, partial $\eta^2 = 0.01$). The growth in knowledge from pre- to post-experimental testing was highest in the active condition (19.77% increase) and lowest in the passive condition (13.92% increase). However, these differences were not statistically significant ($F(2, 88) = 31.43$, $p < 0.68$, partial $\eta^2 = 0.01$; cf. Table 1), which might be due to the fact, that the students' performances varied greatly within and across conditions (cf. Table 1).

3.2. Stress responses during experiments depending on the work setting

3.2.1. Heart rate variability (HRV)

For the LF/HF ratio, we found a significant medium-sized, between-subjects effect of work setting ($F(2, 85) = 6.643$, $p = 0.002$, partial $\eta^2 = 0.135$) and a significant interaction of work setting and the within-subjects variable of time ($F(8.62, 366.54) = 2.249$, $p = 0.02$, partial $\eta^2 = 0.05$; see Fig. 1). In line with Hypothesis 1 (experiments cause different stress responses, depending on

Table 1
Descriptive statistics for the study variables.

Time of measurement	Passive condition	Active condition	Interactive condition
LF/HF ratio			
n	28	40	20
5 min	1.10 ± 0.36	1.07 ± 0.30	1.02 ± 0.28
10 min	1.17 ± 0.36	1.03 ± 0.27	0.86 ± 0.22
15 min	1.29 ± 0.44	1.03 ± 0.23	0.91 ± 0.21
20 min	1.31 ± 0.46	1.14 ± 0.34	0.95 ± 0.29
25 min	1.23 ± 0.35	1.13 ± 0.33	1.01 ± 0.30
30 min	1.39 ± 0.57	1.18 ± 0.30	0.98 ± 0.20
35 min	1.14 ± 0.34	1.14 ± 0.35	1.08 ± 0.29
rmsSD [ms]			
n	28	40	20
5 min	68.34 ± 27.35	87.26 ± 25.74	88.80 ± 26.14
10 min	62.07 ± 28.21	87.51 ± 29.20	105.26 ± 24.00
15 min	54.99 ± 24.71	89.69 ± 28.46	107.61 ± 26.60
20 min	58.53 ± 34.24	82.17 ± 30.05	105.72 ± 33.17
25 min	60.02 ± 31.96	80.16 ± 31.01	99.57 ± 36.61
30 min	59.30 ± 39.44	76.65 ± 29.90	99.65 ± 30.89
35 min	68.57 ± 34.00	77.26 ± 26.38	85.56 ± 29.73
Cortisol [nmol/L]			
n	27	36	18
Before experiment (t0)	12.97 ± 6.34	12.45 ± 4.77	16.84 ± 7.24
During experiment (t0 + 15 min)	10.96 ± 4.73	11.15 ± 5.75	12.41 ± 5.03
After experiment (t0 + 30 min)	10.11 ± 4.57	11.27 ± 5.49	12.61 ± 5.78
VAS [mm]			
n	31	40	20
Before experiment (t0)	11.45 ± 14.95	10.53 ± 9.53	11.10 ± 13.68
During experiment (t0 + 15 min)	19.81 ± 16.76	11.85 ± 11.00	7.45 ± 10.15
After experiment (t0 + 30 min)	15.71 ± 15.78	10.85 ± 12.10	5.20 ± 7.11
Enjoyment			
n	31	39	20
Before experiment	4.10 ± 0.73	4.00 ± 0.74	3.85 ± 1.04
After experiment	3.68 ± 0.88	3.42 ± 1.05	2.42 ± 0.70
Boredom			
n	31	39	20
Before experiment	1.42 ± 0.69	1.46 ± 0.71	1.36 ± 0.93
After experiment	1.35 ± 0.44	1.87 ± 0.85	2.89 ± 1.13
Knowledge [points]			
n	19	30	39
Before experiment	4.71 ± 1.13	4.44 ± 0.81	4.53 ± 0.83
After experiment	5.14 ± 0.87	5.24 ± 1.01	5.29 ± 0.90

Note. Means ± SD for the study variables at the different measurement points.

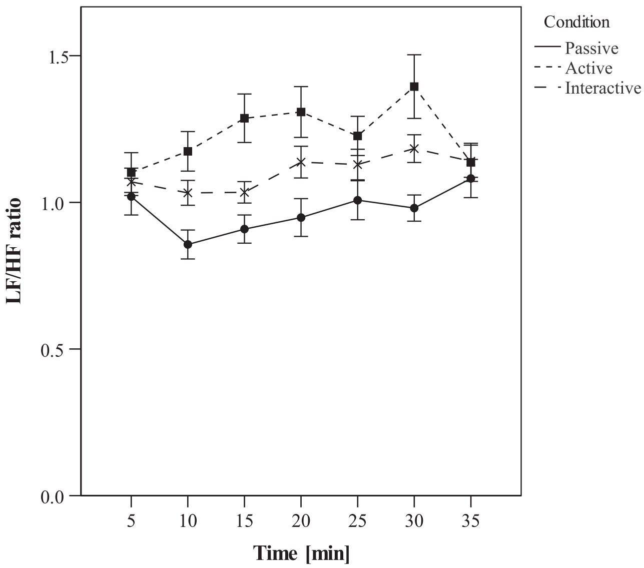

Fig. 1. Students' LF/HF ratio while conducting the experiment. Data represent means ± *SEM* ($n_{passive}$ = 20, n_{active} = 28, $n_{interactive}$ = 40). LF/HF = low frequencies/high frequencies.

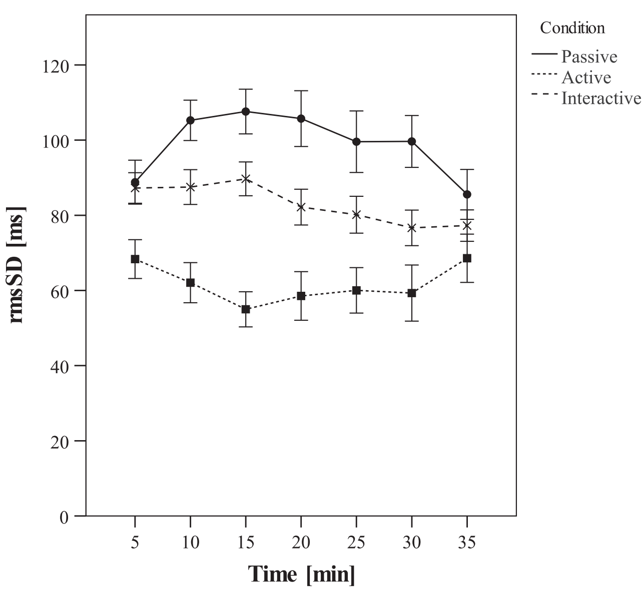

Fig. 2. Students' values of rmsSD while conducting the experiment. Data represent means ± *SEM* ($n_{passive}$ = 20, n_{active} = 28, $n_{interactive}$ = 40). rmsSD = root mean square of successive differences.

the work setting), the students in the passive condition consistently displayed significantly lower LF/HF ratio values, compared to those who performed the task by themselves in the active condition ($p = 0.001$), who showed the highest values. The values displayed by the students in the interactive condition ranged in between and did not differ significantly from the other two conditions. All students started and ended with similar LF/HF ratio values, independent of the work setting (cf. Table 1).

For rmsSD, we found a significant interaction of work setting and the within-subjects variable of time ($F(8.55, 363.25) = 3.57$, $p < 0.001$, partial $\eta^2 = 0.077$). Also as predicted, there was a significant, large between-subjects effect of work setting ($F(2, 85) = 14.233$, $p < 0.001$, partial $\eta^2 = 0.251$; see Fig. 2). The students in the active condition consistently showed significantly lower rmsSD values, compared to the values obtained for the students in the interactive condition ($p = 0.002$). In the interactive condition, the students' rmsSD values stayed constant during the first 15 min, and decreased thereafter. The highest rmsSD values were observed in the passive condition, indicating the lowest level of stress. They were significantly higher compared to the rmsSD values of the students in the active condition ($p = <0.0001$) but did not differ from those of the students in the interactive condition ($p = 0.051$). After conducting the experiment, the rmsSD values were similar across all work settings (cf. Table 1).

3.2.2. Cortisol

In all three conditions, the concentration of salivary cortisol decreased over the course of the experiment, but with a different magnitude ($F(2, 75) = 4.14$, $p = 0.020$, partial $\eta^2 = 0.099$; see Fig. 3, Table 1). As outlined in Hypothesis 1 (experiments cause different stress responses, depending on the work setting), the cortisol concentration of the students in the interactive condition showed considerably less decrease, indicating higher levels of stress, than those of the students who watched the video in the passive condition ($p = 0.014$). Students who conducted the experiment individually in the active condition did not differ significantly from either group. Neither the students' sex, nor the intake of oral contraceptives influenced the cortisol concentration (all p's > 0.09).

3.2.3. Subjective stress perception

In line with Hypothesis 1 (experiments cause different stress responses, depending on the work setting), the students who conducted the experiment individually (active condition) or in groups (interactive condition) showed an increase in their perceived stress during the experiment. In contrast, the stress levels of the students in the passive condition decreased. Accordingly, we found a significant interaction between work setting and the within-subjects variable of time ($F(3.65, 160.67) = 3.28$, $p = 0.016$, partial $\eta^2 = 0.07$). Students working in the active condition felt significantly more stressed while conducting the experiment, as compared to the passive condition ($p = 0.005$). In contrast, the perceived stress of the students working in the interactive condition differed neither significantly from the students in the active condition, nor from the ones in the passive condition (see Fig. 4, Table 1). Overall, the students reported relatively mild stress levels, with the highest mean value during the experiment (active condition: $M = 19.81 \pm 16.73$ *SD*). The students' sex did not modulate the subjective stress levels (all p's > 0.83).

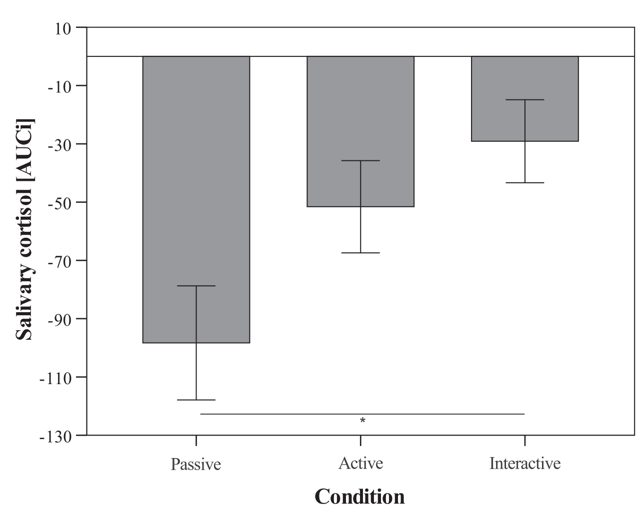

Fig. 3. Students' salivary cortisol change while conducting the experiment ($^{*}p < 0.05$). Data represent means ± *SEM* ($n_{passive}$ = 18, n_{active} = 27, $n_{interactive}$ = 36). AUCi = area under the curve with respect to increase.

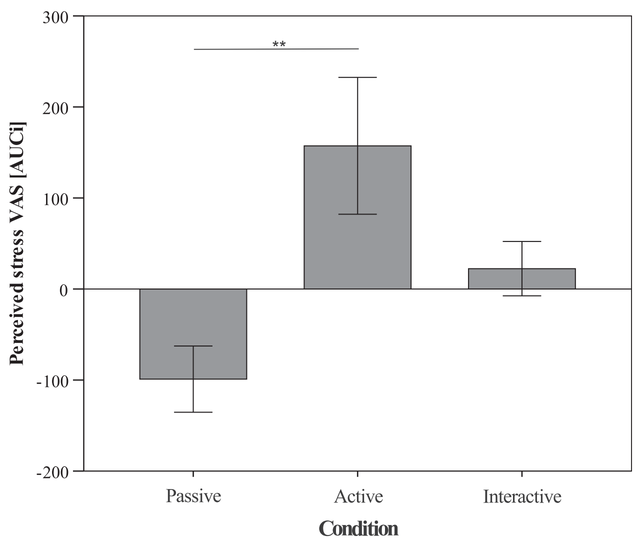

Fig. 4. Students' subjective stress perception (VAS) while conducting the experiment ($^{**}p < 0.01$). Data represent means ± *SEM* ($n_{passive} = 20$, $n_{active} = 31$, $n_{interactive} = 40$). VAS = visual analogue scale; AUCi = area under the curve with respect to increase.

3.2.4. Emotions

As we have stated in Hypothesis 1 (experiments cause different stress responses, depending on the work setting), the work setting had a great effect on the changes in students' enjoyment ($F(2, 87) = 8.13$, $p < 0.001$, partial $\eta^2 = 0.16$) and in boredom ($F(2, 87) = 19.67$, $p < 0.0001$, partial $\eta^2 = 0.31$). Starting off at equally high levels before the experiment ($F(2, 87) = 0.60$, $p = 0.55$), enjoyment in the passive condition dropped sharply, compared to a moderate decrease in the active ($p = 0.003$) and interactive group settings ($p = 0.023$) (see Fig. 5). The latter two settings showed a similar change pattern ($p = 0.59$) (cf. Table 1). Subsequently, repeated-measures ANOVAs for enjoyment were conducted separately for each condition. These analyses identified a strong effect of time, with enjoyment dropping in the passive condition ($F(1, 12) = 15.43$, $p < 0.01$, partial $\eta^2 = 0.56$), the active condition ($F(1, 20) = 7.78$, $p < 0.05$, $\eta^2 = 0.28$), and the interactive condition ($F(1, 26) = 16.22$, $p < 0.0001$, partial $\eta^2 = 0.38$).

Also starting off at similarly low levels before the experiment ($F(2, 87) = 0.11$, $p = 0.89$), boredom showed a steep increase in the passive condition, while the levels almost remained constant among the students in the active condition ($p = 0.001$) and in the interactive condition ($p = 0.04$). Examining the change patterns separately for each condition, subsequent ANOVAs identified a strong increase in boredom for the passive condition ($F(1, 12) = 16.23$, $p < 0.01$, partial $\eta^2 = 0.58$), a moderate increase for the interactive condition ($F(1, 26) = 4.90$, $p < 0.0001$, partial $\eta^2 = 0.38$), and no change for the active condition ($F(1, 20) = 0.16$, $p = 0.69$, partial $\eta^2 = 0.008$) (see Fig. 5, Table 1). The students' sex did not modulate the changes in achievement emotions (all p's > 0.32).

3.3. Associations between emotional and physiological stress measurements

As outlined in hypothesis 2 (different stress parameters are associated with each other), subjective stress levels in the passive condition exhibited a strong positive correlation with the LH/HF ratio ($r = 0.59$, $p = 0.005$), and a negative association with the rmsSD ($r = -0.58$, $p = 0.038$). For the active and the interactive condition, no significant relations could be found, though all patterns pointed in the correct direction. Across all conditions, no significant associations between enjoyment or boredom and stress measurements were found (all p's > 0.086). Examining the remaining associations, a decrease in enjoyment was linked to a strong increase in boredom across all three conditions (all p's < 0.01). Further, the LF/HF ratio exhibited negative associations with the rmsSD levels across all three conditions (all p's < 0.01; see Table 2).

4. Discussion

The current study examined patterns in physiological, subjective and emotional stress reactions, and the associations between these variables, when students conducted a biology experiment under three constructive conditions, which were enriched either with passive, active, or interactive elements, within the framework of CVT. Thereby, we wanted to clarify which characteristics of the work setting, especially if independent from a specific content, influence the students' stress reactions.

4.1. Influence of the work setting on stress responses during experiments

As outlined in Hypothesis 1 (experiments cause different stress responses, depending on the work setting), we found that the work setting has a large impact on the students' subjective, physiological, and emotional stress responses when they conduct a biology experiment in three different constructive settings with inquiry-based elements. In line with the propositions by CVT (Pekrun, 2006; Pekrun et al., 2010), our findings suggest that different combinations of work characteristics, as conceptualized by Chi's taxonomy of work settings enhance differential patterns of stress

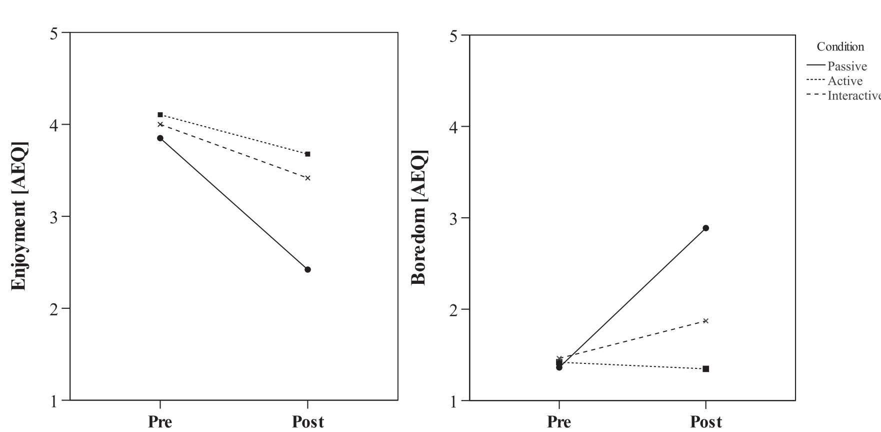

Fig. 5. Students' levels of enjoyment and boredom before and after the experiment ($n_{passive} = 20$, $n_{active} = 31$, $n_{interactive} = 39$).

Table 2
Nonparametric partial correlations between AUCg-values of physiological and psychological stress parameters and emotions controlling for frequency of experiments in biology lessons, biology grade and growth in knowledge from pre- to post-experimental testing.

Measure	2.	3.	4.	5.	6.
Passive condition					
1. LF/HF ratio	−0.72**	−0.04	0.59*	0.24	−0.45
2. rmsSD		−0.20	−0.58*	0.11	0.11
3. Cortisol			−0.18	−0.36	0.49
4. VAS				0.06	−0.34
5. Enjoyment					−0.80*
6. Boredom					
Active condition					
1. LF/HF ratio	−0.78**	−0.38	0.22	−0.05	−0.10
2. rmsSD		0.32	−0.16	0.28	−0.12
3. Cortisol			−0.20	0.15	−0.11
4. VAS				0.02	0.05
5. Enjoyment					−0.78**
6. Boredom					
Interactive condition					
1. LF/HF ratio	−0.72**	0.17	0.12	−0.09	−0.08
2. rmsSD		−0.12	−0.25	−0.02	0.07
3. Cortisol			0.03	0.05	−0.24
4. VAS				−0.08	−0.17
5. Enjoyment					−0.66**
6. Boredom					
Full sample					
1. LF/HF ratio	−0.60**	0.02	−0.01	0.01	−0.21*
2. rmsSD		−0.03	0.02	−0.09	0.28**
3. Cortisol			0.05	0.14	−0.10
4. VAS				0.10	−0.16
5. Enjoyment					−0.58**
6. Boredom					

Notes. * $p < 0.05$; ** $p < 0.01$; LF/HF = low frequencies/high frequencies; rmsSD = root mean square of successive differences; VAS = visual analogue scale; AUCg = Area under the curve with respect to ground.

reactions and emotions. The patterns may be interpreted as follows: If passive features are implemented, students in constructive settings face a low risk of failure, as they do neither have to conduct the experiment themselves nor present their result in the plenum. However, they may suffer from low mental and practical engagement. Thus, subjective and HRV stress responses are thus lowest but come with the cost of high decreases in enjoyment and high increases in boredom (Pekrun & Stephens, 2012; Pekrun et al., 2010).

Conducting experiments individually in constructive settings engages students cognitively, emotionally and behaviorally in the task, which is conducive to maintain enjoyment and prevent a drop in boredom. Thus, changes in activity-related achievement emotions may occur, but to a smaller degree than in constructive settings with passive features (cf. Pekrun, 2006; Pekrun et al., 2010). However, as the complexity of the experimental procedures has to be managed by each student separately, cognitive demands are high. Moreover, as the individual performance is salient to oneself, and to some extent to other students or teachers, the risk of self-attributed failure increases which may be accompanied by subjective, emotional, and/or physiological stress reactions (Lindahl et al., 2005; Pekrun, 2006; Ringeisen, Raufelder, Schnell, & Rohrmann, 2015).

Comparing constructive settings with individual and collaborative features, working on an experiment in a group may be less stressful, as the presence of other students represents potential sources of support (cf. Malecki & Demaray, 2003). In cases wherein the knowledge and engagement of the participating students is high, collaborative experimentation settings may stimulate knowledge exchange, cross-cueing, and error pruning which contribute to better problem-solving capacities, and higher productivity (Gillies, 2004; LoGiudice et al., 2015; Sweller et al., 2011; Zhang et al., 2011). As a consequence, the risk of (self-attributed) failure decreases. Working on an experiment in a group may thus constitute a source of informational and emotional support, which may have positive effects on psychological and academic adjustment (Demaray & Malecki, 2002; Ramsay, Jones, & Barker, 2006). On the cost side, however, collaborative work requires time for coordination, communication, and mutual feedback. Insufficient implementation is often associated with lower group effectiveness, disengagement of individual students, and negative performance evaluations by fellow group members (Arvey & Murphy, 1998; Druskat & Wolff, 1999; Engelmann & Fischer, 2014). These findings imply that collaborative settings are not always interactive and can take on different forms, ranging from active to mainly constructive, as in the current study (Chi, 2009).

Such interpretation is in line with the fact that the smallest cortisol decline, which indicates more stress, was observed among students who worked in the interactive group condition. The cortisol-producing HPA-axis becomes especially activated, when someone perceives a social-evaluative threat from important others—in this case, other students (Dickerson & Kemeny, 2004). However, since the differences in cortisol levels between participants in the interactive and active condition were not significant, and decreasing in both instances, this effect may not be very strong. Conducting an experiment in constructive settings seems to activate the HPA-axis, but is not sufficient to lever out the diurnal cortisol decrease.

In essence, our results suggest that biology experiments with inquiry-based elements in a constructive setting constitute methodically complex and mentally challenging learning situations. They carry the risk of (self-attributed) failure and negative performance evaluations as they require high levels of concentration and scientific reasoning (Hofstein & Lunetta, 2003). Due to these features, experiments with inquiry-based elements may be classified as low to moderate stressors that are accompanied by an increase in subjective and physiological stress parameter and changes in enjoyment and boredom (cf. Engelmann & Fischer, 2014; Pekrun & Stephens, 2012). The intensity of these reactions, however, is affected by the extent to which students are actively engaged in the experiment. If students are highly engaged they may experience elevated stress, but only moderate changes in enjoyment and boredom. Conversely, a lack of practical and mental involvement as in the constructive setting with passive features is accompanied by lower stress, but greater decreases in enjoyment and greater increases in boredom (Pekrun & Stephens, 2012; Pekrun et al., 2010).

4.2. Associations between physiological and psychological stress parameters

All significant associations between physiological, subjective, and emotional stress parameters were in accordance with Hypothesis 2 (different stress parameters are associated with each other). However, a number of predicted associations could not be found.

In line with the results of Looser et al. (2010), subjective stress was related to an increase in the LF/HF ratio, and a decrease in rmsSD in the passive group. Thus, low activity seems be a stressor for students. However, for the students in the active condition, inactivity may be primarily stress-inducing, as it hinders the mastery of the experiment, which leads to worries about failure. In contrast, for the passive condition, inactivity may be rather emotion-inducing, as previous research found mental and practical disengagement to go along with a strong increase in boredom (cf. Pekrun & Stephens, 2012; Pekrun et al., 2010).

In contrast to Spangler et al. (2002) and Lindahl et al. (2005), who found associations between cortisol levels, subjective stress, and negative affective states in academic settings, the current study could not replicate such findings. We observed low to moderate stress levels and a cortisol decline in all three conditions which were unrelated to achievement emotions. Campbell and Ehlert (2012) suggest that non-significant associations may be attributed to situational features, if methodological and inter-individual factors are controlled for as it was the case in our randomized-field experiment with equal assessments across conditions. In line with Campbell and Ehlert (2012), controllability, uncertainty and incentive value may have been at mild to moderate levels across all three conditions, which often results in non-significant associations between different stress parameters (Acee et al., 2010; Cassady & Johnson, 2002; Pekrun et al., 2011; Petermann & Winkel, 2007).

However, pointing to the ecological validity of the data, a decrease in enjoyment was linked to a strong increase in boredom (Pekrun, 2006; Pekrun & Stephens, 2012; Pekrun et al., 2010). In addition, the LF/HF ratio exhibited strong negative associations with the rmsSD levels across all three conditions (Petrowski, Herold, Joraschky, Mück-Weymann, & Siepmann, 2010).

4.3. Strengths and limitations

The present study has a number of strengths. It is the first to investigate whether conducting an experiment is a source of stress for school students, and how different work settings may modulate their stress responses. To minimize the effect of potentially confounding demographic variables, but still implement a real-life laboratory setting, we conducted an experimental field study and randomly allocated the students to each work condition. As studies on learning-related stress typically rely on self-report data, we combined the measurements of physiological, subjective, and emotional stress parameters, which allowed us to assess the stress phenomenon more reliably and investigate if patterns concur (Campbell & Ehlert, 2012; Jennings & Gianaros, 2007).

Besides these strengths, the current study is subject to some methodological limitations. First, our sample size was relatively small, which increases the risk of detecting unusual patterns in temporal change and associations between the various stress indicators. Keeping in mind that the relationships between the various stress parameters may be modulated by situational characteristics (e.g., the experimental task, quality of relationships with others present), methodological (e.g., varying experimental designs and assessments), inter-individual factors (e.g., ability and knowledge of fellow group members), and the overall stress level, linear and non-linear associations patterns may be possible, which may vary in strength by work condition (for an overview, see e.g., Bradley & Lang, 2007; Uchino, Smith, Holt-Lunstad, Campo, & Reblin, 2007). In particular, as participants may vary in their responsiveness of ANS and HPA regulation, psychophysiological correspondence may be biased (Campbell & Ehlert, 2012; Papousek & Schulter, 2001). Future studies should replicate the study with larger samples to estimate the associations between the various stress parameters with greater precision. However, as the concurrent assessment of physiological, subjective, and emotional parameters in real-life learning and academic settings is complex and costly, our sample size can be considered adequate and even bigger than those of similar studies (Jennings & Gianaros, 2007; Spangler et al., 2002).

Second, we investigated students' stress and emotions during variations of partial inquiry-based constructive experimental settings which were adapted to the competence level of older high school students with previous experience in biology. Thus, our findings may not be transferable to other kinds of experiments with different degrees of inquiry, or for other subject domains, and may not apply to younger or older students with different levels of knowledge and experimental competence. However, as we choose an experimental format which is typical for biology classes in high school, we think that our results can be applied to experimental settings containing similar instructive and inquiry-based elements. Third, as we did not measure students' levels of engagement, their perception of situational characteristics, and their appraisals, we encourage future studies to incorporate these variables to examine their interplay with stress responses and emotions across constructive conditions with different experimental features.

4.4. Future research

Based on our findings, future research should focus on what exactly happens during experiments that are performed in groups. Especially the engagement levels of the group members deserve attention. With regard to an optimal group composition it remains unclear which characteristics of the group members, in conjunction with features of the work setting, predict the performance of a group, as well as stress and related emotions of the group members (Demaray & Malecki, 2002; Pluut, Curşeu, & Ilies, 2014; Ramsay et al., 2006).

Moreover, to better understand the relative significance and cumulative impact of stressors in learning and academic settings, future studies should investigate the stress-inducing potential of learning-related activities other than laboratory work, as indicated by physiological, subjective, or emotional measures, and the frequency with which students are exposed to them. Previous studies have shown that stress correlates and emotions differ not only by setting-related activities (e.g., participation in class, doing homework, engaging in group work), but also by subject (e.g., Math, Latin; Goetz, Pekrun, Hall, & Haag, 2006; Goetz et al., 2012; Pekrun et al., 2011). Future research should, thus, assess the stress parameters under question for other subjects (in this study, biology) and setting-related activities (in this study, conducting an experiment).

5. Conclusions

In our study we found a large influence of the work setting on emotions, as well as subjective and physiological stress responses when students conduct an experiment: The strongest stress responses were observed in the active condition, followed by the interactive condition. In the passive condition students showed the lowest stress responses but also the strongest decrease of enjoyment and increase in boredom, indicating passive work settings to be inappropriate for science education.

However, our findings also show that conducting an experiment alone is accompanied by elevated physiological stress responses, as it carries the risk of self-attributed failure. In contrast, conducting an experiment in a group may reduce the stress responses, as others constitute a source of support. At the same time, enjoyment remains on a relatively constant level, comparable to the one of the active work setting. In sum, interactive work settings seems to be most appropriate as they trigger moderate levels of (activating) stress and positive emotions, which may be most beneficial for students' motivation and their attitudes towards science.

If students lack experience and competence in conducting experiments, as they are still rarely used in everyday school life, teachers and experienced peers should serve as assistants to actively engage students in conducting experiments by themselves. This way, the students' enjoyment and their interest in biology topics may be triggered while performance-hindering stress when dealing with novel experimental settings may be reduced. Such combination of achievement emotions helps to build up knowledge of the subject, and may even have beneficial transfer effects on other science domains (Arnold, Kremer, & Mayer, 2014; Dohn et al., 2016; Pekrun & Stephens, 2012).

Funding

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

References

Acee, T. W., Kim, H., Kim, H. J., Kim, J., Hsiang-Ning, R. C., Kim, M., ... Wicker, Frank W. Boredom Research Group. (2010). Academic boredom in under- and over-challenging situations. *Contemporary Educational Psychology, 35*, 17–27. http://dx.doi.org/10.1016/j.cedpsych.2009.08.002.

Agelink, M. W., Malessa, R., Baumann, B., Majewski, T., Akila, F., Zeit, T., & Ziegler, D. (2001). Standardized tests of heart rate variability: Normal ranges obtained from 309 healthy humans, and effects of age, gender, and heart rate. *Clinical Autonomic Research, 11*, 99–108. http://dx.doi.org/10.1007/BF02322053.

Alsop, S., & Watts, M. (2003). Science education and affect. *International Journal of Science Education, 25*, 1043–1047. http://dx.doi.org/10.1080/0950069032000052180.

Arnold, J. C., Kremer, K., & Mayer, J. (2014). Understanding students' experiments – What kind of support do they need in inquiry tasks? *International Journal of Science Education, 36*, 2719–2749. http://dx.doi.org/10.1080/09500693.2014.930209.

Arnold, J., Kremer, K., & Mayer, J. (2013). Wissenschaftliches Denken beim Experimentieren – Kompetenzdiagnose in der Sekundarstufe II (Scientific thinking during experimentation in high school). In D. Krüger, A. Upmeier zu Belzen, P. Schmiemann, A. Möller, & D. Elster (Eds.), *Erkenntnisweg Biologiedidaktik* (pp. 7–20). Kassel: Universitätsdruckerei.

Arvey, R. D., & Murphy, K. R. (1998). Performance evaluation in work settings. *Annual Review of Psychology, 49*, 141–168. http://dx.doi.org/10.1146/annurev.psych.49.1.141.

Baert, S., Casier, A., & De Raedt, R. (2012). The effects of attentional training on physiological stress recovery after induced social threat. *Anxiety Stress Coping, 25*, 365–379. http://dx.doi.org/10.1080/10615806.2011.605122.

Barzel, B., Reinhoffer, B., & Schrenk, M. (2012). Das Experimentieren im Unterricht (Experimenting in the classroom). In W. Rieß, M. Wirtz, B. Barzel, & A. Schulz (Eds.), *Experimentieren im mathematisch-naturwissenschaftlichen Unterricht* (pp. 103–127). Münster: Waxmann.

Bellocchi, A., Quigley, C., & Otrel-Cass, K. (Eds.). (2017). *Exploring emotions, aesthetics and wellbeing in science education research*. New York: Springer.

Berck, K.-H., & Graf, D. (2010). *Biologiedidaktik – Grundlagen und Methoden* (Didactics of biology - Basic principles and methods). Wiebelsheim: Quelle & Meyer.

Bradley, M. M., & Lang, P. J. (2007). Emotion and motivation. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 581–607). Boston: Cambridge University Press.

Campbell, J., & Ehlert, U. (2012). Acute psychosocial stress: Does the emotional stress response correspond with physiological responses? *Psychoneuroendocrinology, 37*, 1111–1134. http://dx.doi.org/10.1016/j.psyneuen.2011.12.010.

Carmona Miranda, K. (2012). *Confirmation and investigation of higher science curiosity in Juarez Middle School students compared to their peers in El Paso, Texas*. <http://digitalcommons.utep.edu/dissertations/AAI1518190> (accessed 08.02.17).

Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology, 27*, 270–295. http://dx.doi.org/10.1006/ceps.2001.1094.

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*, 73–105. http://dx.doi.org/10.1111/j.1756-8765.2008.01005.x.

Chow, A., Eccles, J. S., & Salmela-Aro, K. (2012). Task value profiles across subjects and aspirations to physical and IT-related sciences in the United States and Finland. *Developmental Psychology, 48*, 1612–1628. http://dx.doi.org/10.1037/a0030194.

Danev, S. (2010). *QHRV-lite system guide & user's manual*. Medeia Ltd.

Demaray, M. K., & Malecki, C. K. (2002). The relationship between perceived social support and maladjustment for students at risk. *Psychology in the Schools, 39*, 305–316. http://dx.doi.org/10.1002/pits.10018.

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

Dohn, N. B., Fago, A., Overgaard, J., Madsen, P. T., & Malte, H. (2016). Students' motivation toward laboratory work in physiology teaching. *Advances in Physiology Education, 40*, 313–318. http://dx.doi.org/10.1152/advan.00029.2016.

Dougan, B. K., Hastings, P. D., Granger, D. A., & Usher, B. A. (2001). Adrenocortical activity in at-risk and normally developing adolescents: Individual differences in salivary cortisol basal levels, diurnal variation, and responses to social challenges. *Development and Psychopathology, 13*, 695–719. http://dx.doi.org/10.1017/S0954579401003157.

Druskat, V. U., & Wolff, S. B. (1999). Effects and timing of developmental peer appraisals in self-managing work groups. *Journal of Applied Psychology, 84*, 58. http://dx.doi.org/10.1037//0021-9010.84.1.58.

Engelmann, K., & Fischer, F. (2014). Fostering scientific reasoning: A meta-analysis on intervention studies. In J. L. Polman, E. A. Kyza, D. K. O'Neill, I. Tabak, W. R. Penuel, A. S. Jurow, K. O'Connor, T. Lee, & L. D'Amico (Eds.). *Learning and becoming in practice: The international conference of the learning sciences (ICLS) 2014* (Vol. I, pp. 246–253). Boulder, CO: International Society of the Learning Sciences.

Foley, P., & Kirschbaum, C. (2010). Human hypothalamus-pituitary-adrenal axis responses to acute psychosocial stress in laboratory settings: Psychophysiological biomarkers of health. *Neuroscience and Biobehavioral Reviews, 35*, 91–96. http://dx.doi.org/10.1016/j.neubiorev.2010.01.010.

Gillies, R. M. (2004). The effects of cooperative learning on junior high school students during small group learning. *Learning and Instruction, 14*, 197–213. http://dx.doi.org/10.1016/S0959-4752(03)00068-9.

Goetz, T., Nett, U. E., Martiny, S. E., Hall, N. C., Pekrun, R., Dettmers, S., & Trautwein, U. (2012). Students' emotions during homework: Structures, self-concept antecedents, and achievement outcomes. *Learning and Individual Differences, 22*, 225–234. http://dx.doi.org/10.1016/j.lindif.2011.04.006.

Goetz, T., Pekrun, R., Hall, N., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology, 76*, 289–308. http://dx.doi.org/10.1348/000709905X42860.

Gungor, A., Eryilmaz, A., & Fakioglu, T. (2007). The relationship of freshmen's physics achievement and their related affective characteristics. *Journal of Research in Science Teaching, 44*, 1036–1056. http://dx.doi.org/10.1002/tea.20200.

Hanoch, Y., & Vitouch, O. (2004). When less is more. Information, emotional arousal and the ecological reframing of the Yerkes-Dodson law. *Theory & Psychology, 14*, 427–452. http://dx.doi.org/10.1177/0959354304044918.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., & Hyde, J. S. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology, 106*, 375–389. http://dx.doi.org/10.1177/0956797613518349.

Hellhammer, D. H., Wüst, S., & Kudielka, B. M. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology, 34*, 163–171. http://dx.doi.org/10.1016/j.psyneuen.2008.10.026.

Hjortskov, N., Rissén, D., Blangsted, A. K., Fallentin, N., Lundberg, U., & Søgaard, K. (2004). The effect of mental stress on heart rate variability and blood pressure during computer work. *European Journal of Applied Physiology, 92*, 84–89. http://dx.doi.org/10.1007/s00421-004-1055-z.

Hofstein, A., & Lunetta, V. (2003). The laboratory in science education: Foundations for the twenty-first century. *Science Education, 88*, 28–53. http://dx.doi.org/10.1002/sce.10106.

Isowa, T., Ohira, H., & Murashima, S. (2006). Immune, endocrine and cardiovascular responses to controllable and uncontrollable acute stress. *Biological Psychology, 71*, 202–213. http://dx.doi.org/10.1016/j.biopsycho.2005.04.002.

Itzek-Greulich, H., Flunger, B., Vollmer, C., Nagengast, B., Rehm, M., & Trautwein, U. (2017). Effectiveness of lab-work learning environments in and out of school: A cluster randomized study. *Contemporary Educational Psychology, 48*, 98–115. http://dx.doi.org/10.0.3.248/j.cedpsych.2016.09.005.

Jennings, J. R., & Gianaros, P. J. (2007). Methodology. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 812–834). Boston: Cambridge University Press.

Kärner, T. (2015). *Erwartungswidrige Minderleistung und Belastung im kaufmännischen Unterricht. Analyse pädagogischer, psychologischer und physiologischer Aspekte*. (Underachievement and stress in vocational business education – analysis of pedagogical, psychological, and physiological aspects). Frankfurt a. M.: Lang.

Kemeny, M. E. (2003). The psychobiology of stress. *Current Directions in Psychological Science, 12*, 124–129. http://dx.doi.org/10.1111/1467-8721.01246.

Kirschbaum, C., Tietze, A., Skoluda, N., & Dettenborn, L. (2009). Hair as a retrospective calendar of cortisol production—Increased cortisol incorporation into hair in the third trimester of pregnancy. *Psychoneuroendocrinology, 34*, 32–37. http://dx.doi.org/10.1016/j.psyneuen.2008.08.024.

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist, 46*, 819–834. http://dx.doi.org/10.1037/h0058839.

Lindahl, M., Theorell, T., & Lindblad, F. (2005). Test performance and self-esteem in relation to experienced stress in Swedish sixth and ninth graders – Saliva cortisol levels and psychological reactions to demands. *Acta Paediatrica, 94*, 489–495. http://dx.doi.org/10.1111/j.1651-2227.2005.tb01922.x.

LoGiudice, A. B., Pachai, A. A., & Kim, J. A. (2015). Testing together: When do students learn more through collaborative tests? *Scholarship of Teaching and Learning in Psychology, 1*, 377–389. http://dx.doi.org/10.1037/stl0000041.

Looser, R. R., Metzenthin, P., Helfricht, S., Kudielka, B. M., Loerbroks, A., Thayer, J. F., & Fischer, J. E. (2010). Cortisol is significantly correlated with cardiovascular responses during high levels of stress in critical care personnel. *Psychosomatic Medicine, 72*, 281–289. http://dx.doi.org/10.1097/PSY.0b013e3181d35065.

Luria, R. E. (1975). The validity and reliability of the visual analogue mood scale. *Journal of Psychiatric Research, 12*, 51–57. http://dx.doi.org/10.1016/0022-3956(75)90020-5.

Macher, D., Paechter, M., Papousek, I., & Ruggeri, K. (2012). Statistics anxiety, trait anxiety, learning behavior, and academic performance. *European Journal of Psychology of Education, 27*, 483–498. http://dx.doi.org/10.1007/s10212-011-0090-5.

Malecki, C. K., & Demaray, M. K. (2003). What type of support do they need? Investigating student adjustment as related to emotional, informational, appraisal, and instrumental support. *School Psychology Quarterly, 18*, 231–252. http://dx.doi.org/10.1521/scpq.18.3.231.22576.

Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., & Schwartz, P. J. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use, Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *European Heart Journal, 17*, 354–381. http://dx.doi.org/10.1161/01.cir.93.5.1043.

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *The New England Journal of Medicine, 338*, 171–179. http://dx.doi.org/10.1056/NEJM199801153380307.

Menekse, M., Stump, G., Krause, S., & Chi, M. T. H. (2013). Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education, 102*, 346–374. http://dx.doi.org/10.1002/jee.20021.

Minkley, N., & Kirchner, W. H. (2012). Influence of test tasks with different cognitive demands on salivary cortisol concentrations in school students. *International Journal of Psychophysiology, 86*, 245–250. http://dx.doi.org/10.1016/j.ijpsycho.2012.09.015.

Minkley, N., Westerholt, D. M., & Kirchner, W. H. (2014). Academic self-concept of ability and cortisol reactivity. *Anxiety, Stress, & Coping, 27*, 303–316. http://dx.doi.org/10.1080/10615806.2013.848273.

Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015). The curious case of climate change: Testing a theoretical model of epistemic beliefs, epistemic emotions, and complex learning. *Learning and Instruction, 39*, 168–183. http://dx.doi.org/10.1016/j.learninstruc.2015.06.003.

Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. In S. J. Lopez & C. R. Snyder (Eds.), *Handbook of positive psychology* (pp. 195–206). New York: Oxford University Press.

Noll, A., & Kirschbaum, C. (2006). *Stresskrankheiten vorbeugen und behandeln mit chinesischer Medizin* (Prevention and treatment of stress related diseases with chinese medicine). München: Urban & Fischer.

Osborne, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections. A report to the Nuffield foundation*. London: King's College.

Papousek, I., & Schulter, G. (2001). Associations between EEG asymmetries and electrodermal lability in low vs. high depressive and anxious normal individuals. *International Journal of Psychophysiology, 41*, 105–117. http://dx.doi.org/10.1016/S0167-8760(01)00131-3.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315–341. http://dx.doi.org/10.1007/s10648-006-9029-9.

Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Control-value antecedents and performance consequences of a neglected emotion. *Journal of Educational Psychology, 102*, 531–549. http://dx.doi.org/10.1037/a0019243.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36–48. http://dx.doi.org/10.1016/j.cedpsych.2010.10.002.

Pekrun, R., & Stephens, E. J. (2012). Academic emotions. In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook. Individual differences and cultural and contextual factors. APA handbooks in psychology* (2, pp. 3–31). Washington, DC: American Psychological Association.

Petermann, F., & Winkel, S. (2007). *FLM 7-13 Fragebogen zur Leistungsmotivation für Schüler der 7. bis 13. Klasse* (FLM 7-13 Questionnaire of achievement motivation for students from 7th to 13th grade). Frankfurt a.M., Germany: Harcourt.

Petrowski, K., Herold, U., Joraschky, P., Mück-Weymann, M., & Siepmann, M. (2010). The effects of psychosocial stress on heart rate variability in panic disorder. *German Journal of Psychiatry, 13*, 66–73. http://dx.doi.org/10.1016/j.psyneuen.2009.08.003.

Pluut, H., Curşeu, P. L., & Ilies, R. (2014). Social and study related stressors and resources among university entrants: Effects on well-being and academic performance. *Learning and Individual Differences, 37*, 262–268. http://dx.doi.org/10.1016/j.lindif.2014.11.018.

Priemer, B. (2011). Was ist das Offene beim offenen Experimentieren? (What is open in open ended experiments?) *Zeitschrift für Didaktik der Naturwissenschaften, 17*, 315–337.

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. http://dx.doi.org/10.1016/j.psyneuen.2003.10.002.

Ramsay, S., Jones, E., & Barker, M. (2006). Relationship between adjustment and support types: Young and mature-aged local and international first year university students. *Higher Education, 54*, 247–265. http://dx.doi.org/10.1007/s10734-006-9001-0.

Randler, C., Hummel, E., & Wüst-Ackermann, P. (2013). The influence of perceived disgust on students' motivation and achievement. *International Journal of Science Education, 35*, 2839–2856. http://dx.doi.org/10.1080/09500693.2012.654518.

Rensing, L., Koch, M., Rippe, B., & Rippe, V. (2006). *Mensch im Stress. Psyche, Körper, Moleküle* (Man under stress). München: Spektrum Akademischer Verlag.

Ringeisen, T., Raufelder, D., Schnell, K., & Rohrmann, S. (2015). Validating the proposed structure of relationships among test anxiety and its predictors based on control-value theory: Evidence for gender-specific patterns. *Educational Psychology, 36*, 1826–1844. http://dx.doi.org/10.1080/01443410.2015.1072134.

Rinschede, G. (2007). *Geographiedidaktik* (Didactics of geography). Paderborn: UTB.

Scharfenberg, F.-J. (2010). *Experimentelle Mikrobiologie und Genetik im Biologieunterricht*. (Experimental microbiology and genetics in biology lessons). <http://www.isb.bayern.de/schulartspezifisches/materialien/experimentelle-mikrobiologie-und-genetik-im-biolog> (accessed 07.02.17).

Schilling, T. M., Kölsch, M., Larra, M. F., Zech, C. M., Blumenthal, T. D., Frings, C., & Schächinger, H. (2013). For whom the bell (curve) tolls: Cortisol rapidly affects memory retrieval by an inverted U-shaped dose–response relationship. *Psychoneuroendocrinology, 38*, 1565–1572. http://dx.doi.org/10.1016/j.psyneuen.2013.01.001.

Sinatra, G. M., Broughton, S. H., & Lombardi, D. (2014). Emotions in science education. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 415–436). New York: Routledge.

Sleight, P., & Bernardi, L. (1998). Sympathovagal balance. *Circulation, 98*, 2640. http://dx.doi.org/10.1161/01.CIR.98.23.2640.

Spangler, G., Pekrun, R., Kramer, K., & Hofmann, H. (2002). Students' emotions, physiological reactions, and coping in academic exams. *Anxiety, Stress & Coping, 15*, 413–432. http://dx.doi.org/10.1080/1061580021000056555.

Spering, M., Wagener, D., & Funke, J. (2005). The role of emotions in complex problem-solving. *Cognition and Emotion, 19*, 1252–1261. http://dx.doi.org/10.1080/02699930500304886.

Sweller, J., Ayres, P. L., & Kalyuga, S. (2011). *Cognitive load theory. Explorations in the learning sciences, instructional systems and performance technologies.*. New York: Springer.

Szalma, J. L. (2012). Individual differences in stress, fatigue, and performance. In P. A. Hancock & J. L. Szalma (Eds.), *Performance under stress* (pp. 323–357). Hampshire: Ashgate Publishing.

Uchino, B. N., Smith, T. W., Holt-Lunstad, J., Campo, R., & Reblin, M. (2007). Stress and illness. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 608–632). Boston: Cambridge University Press.

Westermann, J., Demir, A., & Herbst, V. (2004). Determination of cortisol in saliva and serum by a luminescence-enhanced enzyme immunoassay. *Clinical Laboratory, 50*, 11–24.

Wigfield, A., Tonks, S., & Klauda, S. L. (2009). Expectancy-value theory. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 55–75). New York: Routledge.

Wolf, O. T. (2009). Stress and memory in humans: Twelve years of progress? *Brain Research, 1293*, 142–154. http://dx.doi.org/10.1016/j.brainres.2009.04.013.

Zeidner, M. (2007). Test anxiety in educational contexts: Concepts, findings, and future directions. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 165–184). San Diego, CA: Academic Press.

Zhang, L., Ayres, P., & Chan, K. (2011). Examining different types of collaborative learning in a complex computer-based environment: A cognitive load approach. *Computers in Human Behavior, 27*, 94–98. http://dx.doi.org/10.1016/j.chb.2010.03.038.