Comprehensive Psychoneuroendocrinology 10 (2022) 100115

Contents lists available at ScienceDirect

Comprehensive Psychoneuroendocrinology

journal homepage: www.sciencedirect.com/journal/comprehensive-psychoneuroendocrinology

Revisiting the SECPT-G: A template for the group-administered socially evaluated cold-pressor test to robustly induce stress

Benjamin Buttlar [a,*], Helena Dieterle [a], Regan Mandryk [b]

[a] University of Trier, Universitätsring 15, 54296, Trier, Germany
[b] University of Saskatchewan, 110 Science Place, Saskatoon, SK, S7N5C9, Canada

ARTICLE INFO

Keywords:
simultaneous Stress induction
Group setting
Stress response
Hypothalamus-pituitary-adrenal axis
Sympathetic nervous system
Economy

ABSTRACT

The Socially Evaluated Cold-Pressor Test (SECPT (Schwabe et al., 2008) [1]; reliably elicits stress responses. We refined the group-administered version of the SECPT (SECPT-G) aiming to increase its' effectiveness. In Experiment 1 (N = 39), we gathered data from 12 participants simultaneously, employing a stress confederate for each participant. In Experiment 2 (N = 69), we gathered data from six participants simultaneously, employing either six stress confederates (individual-observation) or a single one (group-observation). In Experiment 1, we found that the SECPT-G elicited cortisol responses compared to a control condition; in Experiment 2, we replicated these findings and observed that cortisol responses were similar in the individual- and the group-observation setting. The findings of Experiment 2 were corroborated by people's subjective stress experience. Importantly, both experiments show a similar magnitude of cortisol response, and a greater responder rate than in the regular SECPT or the regular Trier Social Stress Test (TSST). The presented SECPT-G template may thus serve as a reliable and efficient stress induction tool that allows standardization across research groups.

1. Introduction

To gather insights on the human stress response, several paradigms were developed to induce stress. These paradigms helped to understand that people adapt to stress in two waves: The first wave includes immediate reactions to stress and comprises the activation of the sympathetic nervous system (SNS). This is reflected by increased heart rate [2], blood pressure [3], and salivary alpha-amylase concentration (sAA; [4]. Then, a second wave response follows which includes the activation of the hypothalamus-pituitary-adrenal (HPA) axis. This is accompanied by heightened levels of cortisol in blood or saliva [5].

In the first laboratory stress protocol—the Cold-Pressor Test (CPT [6];—people immersed one hand into ice water. This physiological stressor mostly affects the SNS but not the HPA axis [1]. Psycho-social stressors, however, elicit reactions on the SNS and the HPA axis (e.g., Refs. [1,7]. For example, during the Trier Social Stress Test (TSST [7]; stress responses are elicited by a mock job interview, including a free speech and an arithmetic task being observed by a camera and an audience of two or three judges. Meta-analyses show that the effectiveness of psycho-social stressors such as the TSST relies on two components: the social evaluation and the uncontrollability in a situation [8].

Although the TSST creates socially evaluative and uncontrollable situations, it is resource-intensive. The TSST requires the attendance of multiple judges for approximately 15 min who need to be trained extensively. This is because the conversation is unscripted. Therefore, its efficacy depends on the experience of the judges, how well they have been trained, and how the protocol is conducted. These dependencies have made standardization across researchers and laboratories difficult [9].

To increase efficiency and standardization, researchers have combined psycho-social and physiological stressors [10]. Schwabe et al. [1] expanded the CPT by including an observer and a camera monitoring the participant. This Socially Evaluated Cold-Pressor Test (SECPT) creates a socially evaluative and uncontrollable situation similar to the TSST—affecting the SNS and the HPA axis [1]. However, the SECPT is highly standardized because it does not involve free speech; it is also more efficient and less resource-intensive requiring only one observer for 3 min beside an (optionally employed) experimenter [10].

Yet, both paradigms remain labor-intensive: Beside the experimenter, the SECPT requires one [1] and the TSST at least two additional persons [7,9]. To decrease resource constraints, they have been adapted

* Corresponding author.
E-mail address: buttlar@uni-trier.de (B. Buttlar).

https://doi.org/10.1016/j.cpnec.2022.100115
Received 11 January 2022; Accepted 11 January 2022
Available online 22 January 2022

for group settings: The TSST has been administered with up to six [11] and the SECPT with up to 12 participants [12].

While research demonstrates unequivocally that the TSST-G reliably induces stress (e.g. Refs. [11,13], the literature on the SECPT-G shows mixed results: On the one hand, the SECPT-G leads to typical increases of cortisol on average over time [14,15]. On the other hand, reported responder rates are considerably lower after the SECPT-G compared to the regular SECPT in which 60–65% of participants usually classify as responders [10]. For instance, in the study by von Dawans et al. [15] who used the protocol outlined by Minkley et al. [12] only 47% of the participants classified as responders (cortisol increase $\geq$2 nmol/l from baseline to peak).

These mixed results on the SECPT-G may arise because protocols so far [12,14,15] diverge from recommendations of how to conduct the SECPT [10]: In the protocol by Minkley et al. (2014) one experimenter was in charge of the experiment who simultaneously was the observer during the stress induction, which makes it difficult to be neutral and avoid reinforcing feedback; the experimenter did not take notes and cameras were positioned at the corner of the room, which might feel less evaluative; lastly, participants were told how long they would have to immerse their hands into the ice-cold water, which reduces the perceived uncontrollability. In the protocol by Becker et al. [14] no cameras were present and experimenters did not take notes; instead, a competitive situation was created in which participants observed each other while they immersed their hands into the ice-cold water. Additionally participants saw a countdown of how long they would have to immerse their hand into the water.

1.1. *The present investigation*

In two experiments, we aimed to establish and evaluate a refined protocol of the SECPT-G. Here, we were especially interested in the second wave response as previous studies on the SECPT-G did not yield comparable cortisol responses compared to the SECPT. Moreover, we assessed participants first-wave stress response by measuring sAA as previous research failed to show effects on this measure [14]. Lastly, we assessed participants' subjective stress experience.

In Experiment 1, we assessed twelve participants who simultaneously immersed their hands into ice-cold water. To stick as close to the regular SECPT and to the recommendations by Schwabe and Schächinger [10] as possible, each participant was watched by an individual observer. This way, we aimed to show that the SECPT-G can robustly induce stress responses. In Experiment 2, we then tested whether the SECPT-G can be conducted with fewer observers, increasing efficiency and standardization while retaining a high degree of social evaluation. Thus, we assessed six participants simultaneously, who were watched by one observer each (i.e., six observers) or by one observer for all six participants. In both experiments, we included a control condition in which participants immersed their hands into warm water and were neither watched nor recorded.

This research was conducted in accordance with the declaration of Helsinki and approved by the Behavioural Research Ethics Board of the University of Saskatchewan (BEH:#176). Below, we state all measures, manipulations, and exclusions; all materials, data, and analysis scripts are available online on https://osf.io/jckhx/.

2. Experiment 1

2.1. *Methods*

2.1.1. *Participants and design*

The study was conducted in a 2 (condition [stress, control]) x 6 (time of measurement [(−25,-1,+1,+20,+30,+45)]) mixed factors design with repeated measures on the second factor. Participants were randomly assigned to the conditions. The required sample size was determined a priori by using G*Power 3.1 [16], assuming a correlation of $r = 0.5$ for repeated measures, a power of $1\text{-}\beta = 0.95$, and an effect size of $\eta^2 = 0.11$. This effect size was found in the initial SECPT-G protocol by Minkley et al. [12] when assessing its' effects on cortisol increase. This power analysis resulted in a total of $N = 16$ participants for the two-way interaction. To account for the possibility that effect sizes might be smaller in our experiment as well as for the occurrence of outliers and participant non-compliance in a study session, we aimed to recruit a sample of $N = 48$ (i.e., 24 participants per condition). In two sessions, one participant did not attend, respectively. Thus, a member of the research lab acted in all ways as a participant to keep the group sizes equal; this data is not included in the analysis. The final sample thus consisted of $N = 46$ (i.e., 23 participants per condition).

Participants were recruited via online announcement at the University of Saskatchewan (Canada) and compensated with $20 CAD. Exclusion criteria (all self-reported) included obesity (BMI $\geq$ 30 kg/m^2) and underweight (BMI < 18 kg/m^2), excessive smoking, alcohol or drug consumption, mental disorders or major illnesses, as well as currently taking medications with potential effects on the HPA axis (for a review see Kudielka et al. [17]. Participants were also required to refrain from excessive exercise and consumption of meals, caffeine, and alcohol within three hours prior to the study. These criteria were stated in the announcement and emphasized in the confirmation e-mail that was sent to the participants after they signed up; however, seven participants did not meet the exclusion criteria: Two participants stated to meet criteria of clinical depression, one of obsessive-compulsive disorder, one of post-concussion syndrome and three participants had a BMI of 30 kg/m^2 or higher. These seven participants were therefore excluded from statistical analysis. The sample used for analyses thus composed of $N = 39$ (20 stress condition; 19 control condition) healthy adolescents and adults ($M_{\text{age}} = 22.05$, age range = 17–33, $M_{BMI} = 22.53$ kg/m^2, BMI range = 18–29 kg/m^2), 22 female and 17 male. This final sample suffices to find small to medium-sized effects of $\eta^2 = 0.04$ with a power of $1\text{-}\beta = 0.95$ for the hypothesized interaction [16].

2.1.2. *Procedure and materials*

The 90-min experimental sessions took place between 15:00 and 19:30 hours to control for diurnal variations of cortisol [17]. Prior to the experiment, all participants were informed about the procedure, the fact that they could be videotaped during the experiment (only participants in the stress condition were actually filmed), and their right to withdraw from participation at any time; then they provided written consent. After completion of the experiment, participants were debriefed.

The laboratory was set up with three rows of tables that were separated by two corridors (see Fig. 1). Participants were seated on both sides of the corridor; this enabled the experimenters to access the participants during the SECPT-G. To highlight the socially evaluative component, cameras were attached visibly on the corridor-facing side of the computer display.

The experiment and timing were controlled using online software that combined tasks, instructions, and questionnaires. This enabled participants to complete the experiment mostly by themselves guaranteeing control across participants and experimental sessions. Timestamps allowed the stress measurements to occur in exact time intervals for all participants. Thereby, we avoided unsystematic variation due to interaction with experimenters and time dissimilarities due to manual instructions. As in previous studies [10], we measured our dependent variables at six time points (−25, −1, +1, +20, +30, +45; indicated relative to stressor onset (−) and offset (+) in minutes).

Between stress measures, all participants completed questionnaires, including demographics, which assessed sociodemographic variables (age, gender, education), anthropometric data (height, weight), and exclusion criteria. To prevent participants from being interrupted by the stress measurement time points, we kept tasks short and presented casual and relaxing filler games—clicking on colorful dots, an infinite runner game, whack-a-mole—to those participants who finished a task before the next measurement. A visualization of the experimental

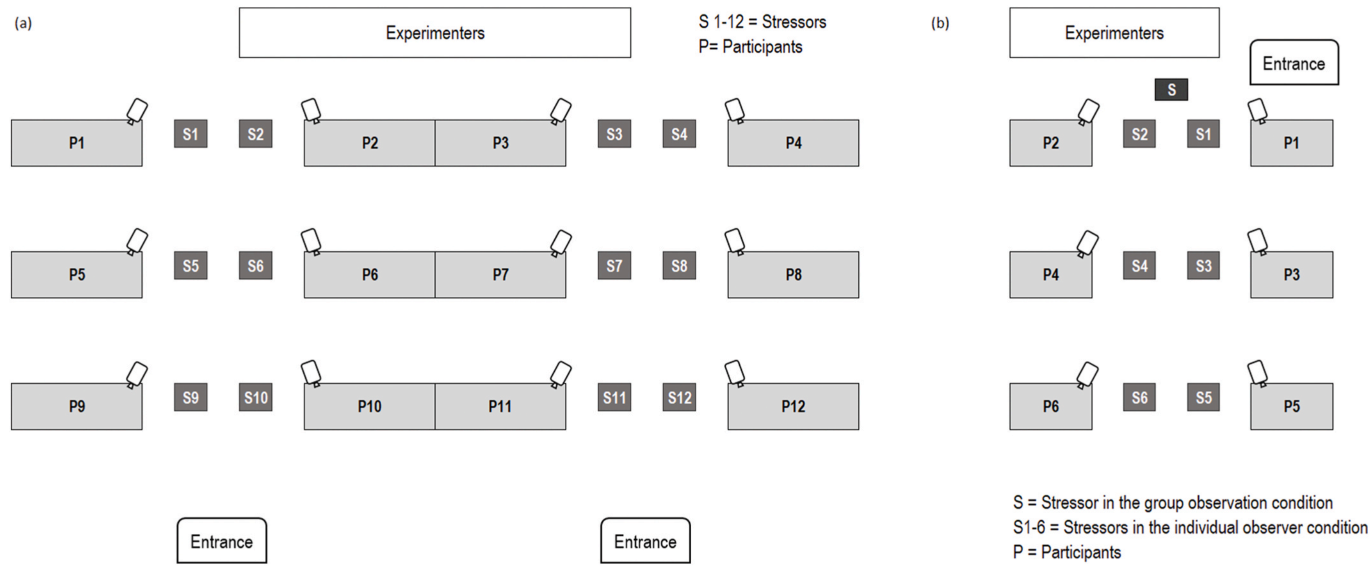

Fig. 1. (a) Depiction of the laboratory setting in Experiment 1 including the position of participants (P), stressors (S1-12), and experimenters (front). (b) Depiction of the laboratory in Experiment 2 including the position of participants (P), stressor in the group-observation condition (S), stressors in the individual-observation condition (S1-6), and experimenters (front).

procedure can be found on the OSF (https://osf.io/9ycdf/).

2.1.2.1. Individual-observation SECPT-G protocol. To correspond to the single-participant SECPT protocol provided by Schwabe et al. [1]; one stress confederate was responsible for one randomly-assigned participant each. Stress confederates were trained to avoid any reinforcement or interaction, and to maintain a reserved attitude and neutral facial expression. They were mixed-sex and wore a white lab coat with a name badge. Additionally, they were equipped with a water basin filled with ice water (0–2 °C), a thermometer, a water pump to circulate water in the basin, a stopwatch, and a clipboard with an attached observation sheet.

Prior to the stress manipulation (i.e., after the second time of measurement), participants were instructed to prepare their non-dominant arm and lay it on the desk. Two regular experimenters then opened the doors and the twelve stress confederates entered who each observed one of the participants (see Fig. 1 for layout). Without speaking, the stress confederates took a position in front of their assigned participant and put the water basin with ice water onto the desk, plugged in the water pump, and started the recording camera (with the preview video shown on the display). After the agreed signal of the words “Start the session now”, they started to recite a script simultaneously, instructing the participants to immerse their hand up to and including the wrist into the ice water and to keep it there as long as possible while looking into the camera. Participants were not told how long they would have to immerse their hand into the cold water. In addition, the stress confederates closely observed the participants and completed an observation sheet. If the participants took their hand out of the water, the stress confederates asked them to immerse it again; but even if they did not follow these orders, the stress confederates continued their observation for the full-time period of 3 min. On average, participants kept their hand in the water for 2.2 min ($SD = 0.88$). Those who kept their hand in the water for 3 min were informed at that point that they are allowed to remove their hand. The stress confederates then took the equipment they brought with them and left the laboratory. The experimenters then instructed the participants to dry their hands and asked them to continue with the experiment.

2.1.2.2. Control protocol. The control condition followed a similar procedure: The regular experimenters (of which there were four, each responsible for three participants) supplied a water basin filled with warm water (35–37 °C). There were no cameras or stress confederates present who monitored the participants. At the beginning of the protocol, participants were informed that they were assigned to the control condition, then asked to place their hand including the wrist into the water for 3 min, to keep it inside the water for the full-time period, and to remove it once the time was over.

2.1.2.3. Measures

We assessed the stress response via salivary cortisol, sAA, and subjective stress ratings. We examined stress levels six times during the experimental procedure. To collect saliva, we used a commercially available sampling device (Salivette; Sarstedt, Germany). The samples were stored at −20 °C and analyzed at the University of Trier (Germany). To explain to participants how to provide the samples, a subtitled video was presented at the first time of measurement, showing how to insert, chew for 60-s, and deposit the salivettes in the labeled tubes. Subsequent saliva samples were gathered by showing condensed explanations in a Graphics Interchange Format (GIF) alongside a 60-s timer.

Subjective stress was assessed via visual analogue scales (VAS; all α between 0.73 and 0.83; e.g., “How tense are you at this moment?”) consisting of eight items on a scale from 0 to 100. The items were worded as followed: “How are you currently feeling?”; “How afraid do you feel at the moment?”; “How would you estimate your current physical unease”; “How strong is your need to leave this situation?”; “How tense do you feel at the moment?”; “How much are you in control of the situation?”; “How much do you feel stressed out?”; “How strong is your desire to have a familiar person by your side?” [18]. Labels were specific to the question (e.g., “Not at all” and “Very strong” for “How strong is your need to leave this situation?”). Exploratorily, we assessed by two further questionnaires to assess people emotions over time: The Discrete Emotions Questionnaire (DEQ [19]; was assessed via 32 items on a scale ranging from 1 (not at all) to 7 (an extreme amount), was combined into scores for positive (all α between 0.66 and 0.74) and negative (all α between 0.66 and 0.95) emotions; the State-Trait Anxiety Inventory (STAI [20]; was captured via 20 items (all α between 0.90 and 0.93; e.g., “I feel strained”, “I feel comfortable”) on a scale ranging from 1 (strongly disagree) to 4 (strongly agree).

Moreover—though not relevant for our hypotheses—participants filled out the Revised Self-Consciousness Scale [21], the Self Description Questionnaire [22], the Basic Psychological Need Satisfaction Scale [23], the Brief COPE [24], the Intrinsic Motivation Inventory [25], the Flow State Scale [26], the Recovery Experience Questionnaire [27], the trait scale of the STAI [20], and the New General Self-Efficacy Scale [28].

2.2. Results of experiment 1

2.2.1. Cortisol

To test the effectiveness of the SECPT-G in eliciting HPA axis activation, we first classified participants as cortisol responders and non-responders. Our data revealed a responder rate of 80 (75%) percent in the SECPT-G condition and zero percent in the control condition (see Table 1) based on a 2.0 nmol/l baseline-to-peak increase criterion.

We then tested whether the SECPT-G would increase cortisol levels in a 2 × 6 ANOVA with the full sample (i.e., including responders and non-responders). This analysis yielded significant main effects of Condition ($F(1, 37) = 38.25$, $p < .001$, $\eta_p^2 = 0.51$, 95% CI [0.26; 0.65]) and Time (F (5,185) = 17.21, $\varepsilon = .37$, $p < .001$, $\eta_p^2 = 0.32$, 95% CI [0.20; 0.40]). In line with our prediction, these main effects were qualified by a significant interaction ($F(5, 185) = 22.09$, $\varepsilon = 0.37$, $p <.001$, $\eta_p^2 = 0.37$, 95% CI [0.25; 0.45]). Further examination by pairwise comparisons for separate times of measurement showed significant group differences at +20, +30 and + 45 (all p's < 0.001)—which is in line with our hypotheses (see Fig. 2a).[1]

Next, we evaluated the area under the curve values with respect to the ground (AUCg) and with respect to increase (AUCi) [30] via two one-way ANOVA. Using these variables yielded significant main effects for AUCg (F (1, 37) = 37.09, $p < .001$, $\eta_p^2 = 0.50$, 95% CI [0.26; 0.65]) and AUCi (F (1, 37) = 28.00, $p <.001$, $\eta_p^2 = 0.43$, 95% CI [0.18; 0.59]), indicating a cortisol response following the SECPT-G but not in the control condition (see Fig. 2b).

2.2.2. Alpha-amylase

We analyzed sAA data via a 2 × 6 ANOVA predicting that sAA levels—as a marker of the first wave stress response—increases directly following the SECPT-G, but not in the control condition. This analysis showed no significant main effect of Condition (F (1, 37) = 1.06, p = .311); however, the main effect of Time ($F(5, 185) = 9.61$, $\varepsilon = 0.46$, $p <.001$, $\eta_p^2 = 0.21$, 95% CI [0.09; 0.29]) and the two-way interaction were significant (F (5, 185) = 3.73, $\varepsilon = 0.46$, $p < .023$, $\eta^2 = 0.09$, 95% CI [0.01; 0.16]). Yet, when comparing the stress and the control condition, via pairwise comparisons there were no significant differences (all ps > .08) between the conditions at the respective times of measurement (see Fig. 3a).

Table 1
Responder rates by Condition for a liberal and a conservative criterion (2 vs. 2.5 nmol/l baseline-to-peak increase in free salivary cortisol) and response sizes by Condition (absolute and relative baseline-to-peak increase in free salivary cortisol) for Study 1 and 2.

	Responder rate		Response size	
	Response criterion		Increase	
	2 nmol/l	2.5 nmol/l	Absolute	Relative
Study 1				
Individual-Observation	17/20 (80%)	15/20 (75%)	7.70 nmol/l	316%
Control	0/19 (0%)	0/19 (0%)	−0.11 nmol/l	−3%
Study 2				
Group-Observation	19/23 (83%)	18/23 (78%)	10.14 nmol/l	434%
Individual-Observation	16/23 (70%)	16/23 (70%)	7.37 nmol/l	259%
Control	3/23 (13%)	2/23 (9%)	−0.36 nmol/l	−11%

Note. Baseline is defined as time point −25; peak is defined as maximum of time points +1, +20, +30 and + 45.

2.2.3. Subjective stress

We tested subjective stress perception by conducting a 2-way ANOVA using the mean score of all VAS items. The main effects of Condition and Time failed to reach significance (Condition: $F(1, 37) = 1.65$, $p = .207$; Time: $F(5, 185) = 1.98$, $\varepsilon = 0.53$, $p = .396$), as well as the predicted interaction ($F(5, 185) = 0.61$, $\varepsilon = 0.53$, $p = .587$). Eight ANOVAs analyzing all VAS items separately did not generate further evidence for our hypothesis that the stress condition affected participants subjective stress experience (main effects Condition: all $Fs(1, 37) < 2.72$, all $ps > .108$; main effect Time: all $Fs(5, 185) < 5.52$, all $ps > .002$; interaction Condition x Time: all $Fs(5, 185) < 1.45$, all $ps > .226$; all analyses were adjusted for sphericity violations via Greenhouse-Geisser correction). This indicates that subjective stress did not increase after the SECPT-G (see Fig. 3b). Yet, Fig. 3b shows that the descriptive statistics do trend in the expected direction. Notably, ratings on the DEQ and the STAI were also unaffected by the stress manipulation.[2]

3. Experiment 2

Building upon the results of Experiment 1, we wanted to assess whether we could elicit similar stress responses with only one experimenter who observed the whole group when keeping the other changes to previous SECPT-G protocols, i.e., the stress confederate takes notes, making cameras highly salient, and not informing participants about the duration of the stressor [10]. This would allow researchers to conduct the SECPT-G more efficiently by reducing the number of people involved and the necessary coordination. Thus, in Experiment 2, we used a group-observation SECPT-G protocol in which one stress confederate watched the whole group, and compared this approach to a condition applying the individual-observation protocol used in Experiment 1, and a control group.

[1] To correct for violations of sphericity indicated by Mauchly's test, we adjusted via Greenhouse-Geisser ($\varepsilon < .75$) or Huynh-Feldt ($\varepsilon > 0.75$) correction [29], and the according ε was stated. All pairwise comparisons were bonferroni adjusted. Both was done for all of our analyses.

[2] We also conducted a one-way ANOVA to test whether the SECPT-G lead to an increase of state anxiety assessed by the STAI. This analysis resulted in non-significant main effects of Time (F (5, 185) = 0.82, $\varepsilon = 0.59$, $p = .48$) and Condition (F (1, 37) = 3.87, $p = .057$, as well as a non-significant interaction of the factors Condition and Time (F (5, 185) = 2.19, $\varepsilon = 0.59$, $p = .095$. Although pairwise comparisons showed a significant group difference at time point +1 (p = .01) and descriptive statistics are in line with the predicted response pattern, these results have to interpreted cautiously due to the missing interaction effect. Lastly, we tested the hypothesis that stress exposure would influence positive and negative affect, and we applied an ANOVA on ratings of the DEQ for combined values of positive and negative emotions each. Analysis did not reveal statistical significant effects for positive emotions (main effect Condition: F (1, 37) = 0.54, $p = .468$; main effect Time: F (5, 185) = 1.68, $\varepsilon = 0.62$, p = .320; interaction of Condition and Time: F (5, 185) = 2.00, $\varepsilon = 0.62$, $p = .116$), nor for negative emotions (main effect Condition:$F(1, 37) = 1.29$, $p = .263$; main effect Time: F (5,185) = 1.07, $\varepsilon = 0.50$, $p = .360$; interaction of Condition and Time: F (5, 185) = 0.22, $\varepsilon = 0.50$, $p = .852$). To summarize, we could not obtain empirical evidence that emotions following the stress and control manipulation differed from each other.

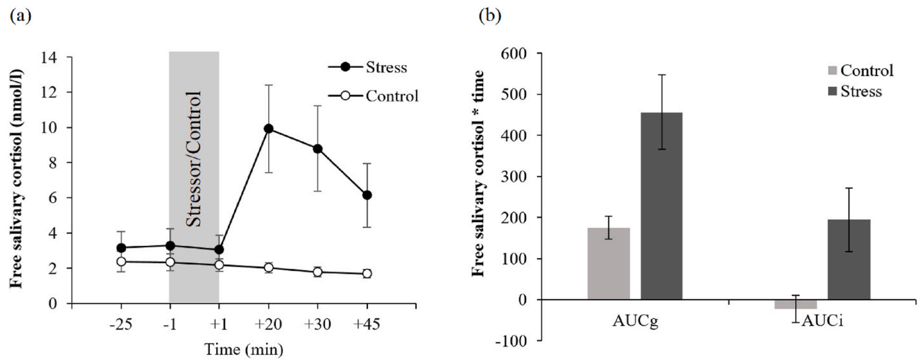

Fig. 2. (a) Concentration of free salivary cortisol in response to the stress and control manipulation sampled at six time points over the course of the experiment and as a function of the experimental conditions. (b) area under the curve with respect to the ground (AUCg) and area under the curve with respect to increase (AUCi). Error bars denote 95% confidence intervals.

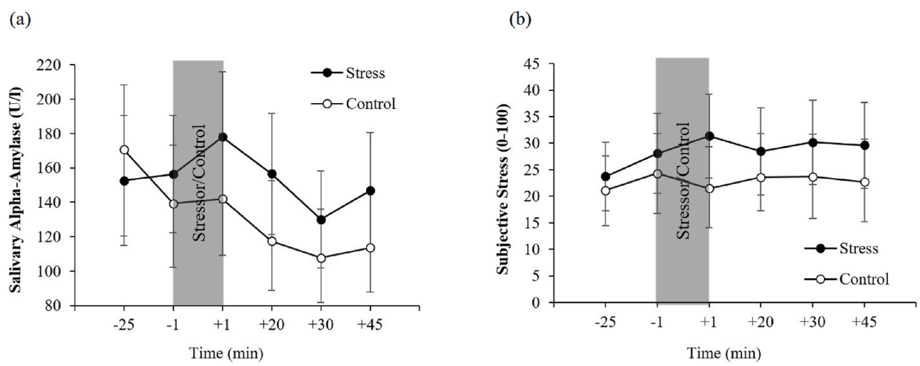

Fig. 3. Concentration of free salivary alpha-amylase (a) and subjective stress reports on the mean of the visual analogue scales (b) in response to the stress and control manipulation sampled at six time points over the course of the experiment and as a function of the experimental conditions.

3.1. Methods

3.1.1. Design and participants

The experiment was conducted in a 3 (condition [individual-observation, group-observation, control]) x 6 (time of measurement [(−25, 1,+1,+20,+30,+45)] mixed factors design with repeated measures on the second factor. We aimed to sample the same number of participants per conditions as in Experiment 1. Participants were recruited as in Experiment 1, applying the same exclusion criteria. The final sample was composed of N = 69 healthy adolescents and adults ($M_{age} = 23.41$, age-range = 17–37, $M_{BMI} = 23.25$ kg/m^2, BMI range = 18.37–29.76 kg/m^2), 27 female, 41 male, 1 non-binary. This sample allows finding an interaction effect of $\eta^2 = 0.03$ with a power of $1\text{-}\beta = 0.95$ assuming a correlation of $r = 0.5$ as in Experiment 1 [16]. Assignment to the conditions was done randomly, adding up to $N = 23$ participants per condition. Participants were compensated with $20 CAD.

3.1.2. Procedure and materials

The 90 min sessions took place between 14:00 h and 18:30. Participants gave informed consent and were debriefed as in Experiment 1.

The laboratory was set up similarly to Study 1, consisting of three rows of tables that were separated by one corridor (see Fig. 1), basically reproducing half the room of Study 1. Cameras were attached visibly on top of the laptop displays. Again, the study was run using online software that combined all tasks, instructions, and questionnaires. Participants changed locations after the stressor, which was done to schedule overlapping sessions within the window of afternoon time available.

In Experiment 2, the time filler was realized by the task of manually completing a sudoku riddle. Moreover, participants were instructed to complete a spatial working memory task and to play a stop-signal game. A depiction of the experimental procedure can be retrieved on the OSF (https://osf.io/brwqt/).

3.1.2.1. Individual-observation SECPT-G protocol. The individual-observation protocol was realized following the same protocol as in Experiment 1: Each stress confederate was randomly assigned to one participant although only 6 instead of 12 participants took part in the experiment simultaneously.

3.1.2.2. Group-observation SECPT-G protocol. In the group-observation protocol, only one person served as a stress confederate. He entered the room and immediately took position in front of the participants in the middle of the corridor and observed the whole group (see Fig. 1). Two main experimenters were responsible to set up the water basins, plug in the water pumps, and start the cameras. Then, the stressor started reciting the same script as in the individual-observation protocol. During the 3 min in which participants immersed their arm into the ice water, the stressor walked up and down the corridor. He observed participants closely and continuously took notes. If participants took their hand out of the water, the stressor asked them to immerse it again.

3.1.2.3. Control protocol. The control protocol was administered in the same way as in Experiment 1, however, with only 6 participants present simultaneously.

3.1.2.4. Measures. We used the same devices and explanations for measuring salivary cortisol and sAA as in Experiment 1. Subjective stress ratings, however, were only assessed via the 8-item VAS (all α between 0.78 and 0.86; [18]. Notably, we also assessed the STAI Trait [20], the Basic Psychological Need Satisfaction Scale [23], and the Revised

Self-Consciousness Scale [21] which are not of concern during the current investigation.

3.2. Results of experiment 2

3.2.1. Cortisol

We again used the 2.0 nmol/l baseline-to-peak increase criterion [15] to examine activation of the HPA axis. Again, we found higher response rates in the SECPT-G conditions compared to the control condition: In the individual-observation condition, 69.6% classified as responders being exceeded by the group-observation condition with 82.6% responders; in the control condition, only 8.7% were classified as responders (see Table 1). A follow-up chi-square test (using only the stress conditions) indicated that the responder rates in the individual and the group-observation condition did not differ χ^2 (1) = 1.08, p = .300.

We then conducted a 3 × 6 ANOVA on cortisol levels. The ANOVA revealed significant main effects of Condition, $F(2, 66) = 6.27$, $p = .003$, $\eta_p^2 = 0.16$, 95% CI [0.02; 0.30], and Time (F (5, 330) = 23.82, $\varepsilon = 0.29$, $p <. 001$, $\eta_p^2 = 0.27$, 95% CI [0.18; 0.33]). As predicted, we found a significant interaction, $F(10, 330) = 11.66$ $\varepsilon = 0.27$, $p <. 001$, $\eta_p^2 = 0.26$, 95% CI [0.16; 0.32] (see Fig. 4a). Pairwise comparisons showed significant group differences between both SECPT-G conditions and the control condition at +20, +30 (all $ps < .011$) and between the control condition and the group-observation condition at + 45 ($p = .001$). Importantly, there were no significant differences between the group-observation and the individual-observation condition at any point in time (all $ps > .391$).

Two one-way ANOVAs using AUCg and AUCi as dependent variables revealed significant main effects (AUCg: F (2, 66) = 4.83, $p = .011$, $\eta_p^2 = 0.13$, 95% CI [0.01; 0.27]; AUCi: F (2, 66) = 10.97, $p < .001$, $\eta_p^2 = 0.25$, 95% CI [0.08; 0.39]). Pairwise comparisons demonstrated that for AUCi, both stress protocols differed from the control condition; however, for AUCg only the group-observation protocol differed from the control condition; the stress conditions did not differ for AUCi and AUCg (see Fig. 4b).

In order to shed more light on whether the group- and the individual-observation protocol differ from each other, we calculated a Bayesian 2 (condition [individual-observation, group-observation]) x 6 (time of measurement [(−25, 1,+1,+20,+30,+45)] mixed ANOVA. This analysis indicated that the best way to represent the data comprises a model that included only the main effect time of measurement. The Bayes factor (BF_{10}) was $4.07*10^{20}$ which indicates decisive evidence in favour of the best model when compared to the null model. The BF_{01} provided strong evidence against adding the main and interaction effects to the best model as it equalled 19.52. This suggests that it is 19.52 more likely that the data is best represented by the model including only time as the predictor, compared to the model including all main and interaction effects.

3.2.2. Alpha-amylase

We tested whether sAA as a marker of autonomic stress response rises as a direct response to the SECPT-G protocols. Therefore, we conducted a 3 × 6 ANOVA, which did not yield a significant main effect of Condition, $F(2, 66) = 0.16$, $p = .850$. Yet, it revealed a significant main effect of Time ($F(5, 330) = 5.91$, $\varepsilon = 0.56$, $p = .001$, $\eta_p^2 = 0.08$, 95% CI [0.03; 0.13]) and a significant two-way interaction (F (10, 330) = 2.20, $\varepsilon = 0.56$, $p = .050$, $\eta^2 = 0.06$, 95% CI [0.002; 0.09]). Pairwise comparisons indicate that the conditions did not differ looking at separate time points (see Fig. 5a).

3.2.3. Subjective stress

We assessed participants' subjective stress levels using the VAS mean in a 3 x 6 ANOVA. The main effect of Condition was not significant, $F(2, 66) = 2.96$, $p = .059$, but there was a significant main effect of Time, $F(5, 330) = 17.00$, $\varepsilon = 0.48$, $p < .001$, $\eta^2 = 0.21$, 95% CI [0.12; 0.27]), as well as a significant interaction, $F(10, 330) = 6.63$, $\varepsilon = 0.48$, $p < .001$, $\eta^2 = 0.17$, 95% CI [0.08; 0.22]). As expected, subjective stress ratings were higher directly after the stressor at +1 (see Fig. 5b) as participants in both stress conditions reported higher levels of subjective stress compared to the control condition (both $ps < .032$) while they did not differ from each other ($p = 1.00$).

4. General discussion

To examine how humans respond to and cope with stress, several protocols were established to induce stress in the laboratory. While group-administrated stress protocols showed promise to reduce the resource constraints, literature is mixed on whether the SECPT-G robustly evokes cortisol responses [12,14,15]. Thus, we refined the SECPT-G according to recommendations of how to induce stress via the SECPT [10]: We made the social evaluation more salient by placing cameras prominently, and the stress confederate(s) continuously took notes; we also highlighted the uncontrollability of the situation by not disclosing how long participants are asked to immerse their hand into the ice water.

In two experiments, the refined SECPT-G reliably lead to cortisol responses that exceed response rates found in studies using single-participant and group-administered SECPTs. This was the case for an individual-observation setting (Experiment 1: 80% and Experiment 2: 69.6%) in which each participant was assigned to one stress confederate and a group-observation setting (Experiment 2: 82.6%) in which only one stress confederate observed the whole group. For comparison, in a previous experiment using a group-administered SECPT, only 47% of

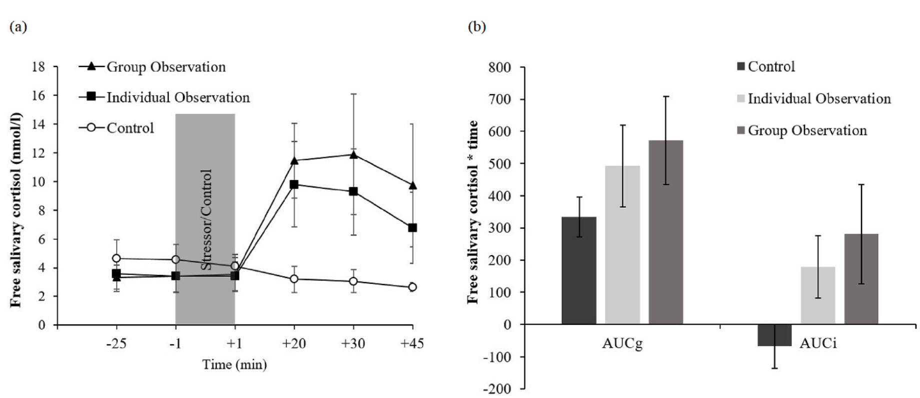

Fig. 4. (a) Concentration of free salivary cortisol in response to the stress and control manipulations sampled at six time points over the course of the experiment and as a function of the conditions. (b) area under the curve with respect to the ground (AUCg) and area under the curve with respect to increase (AUCi). Error bars denote 95% confidence intervals.

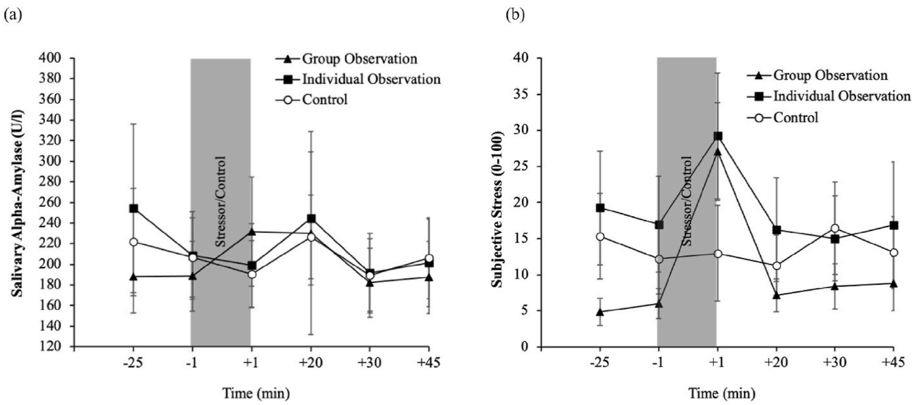

Fig. 5. Concentration of free salivary alpha-amylase (a) and subjective stress reports on the mean of the visual analogue scales (b) in response to the stress and control manipulation sampled at six time points over the course of the experiment and as a function of the experimental conditions.

participants were classified as responders using the same criterion (i.e., cortisol increase ≥2 nmol/l; [15]. Importantly, using this criterion (or a more conservative criterion of cortisol increase ≥2.5 nmol/l), the SECPT-G even shows higher responder rates compared to single SECPT protocols where responder rates range from 60 to 65% [10].

The SECPT-G even compares favorably to TSST protocols which produce higher responder rates than the SECPT at roughly 75% [10]. Indeed, when looking at the cortisol increase and effect sizes of the SECPT-G, both experiments are in line with or exceed the effects of typical cortisol responses found following TSSTs. Across experiments, we found a 2.5-fold–4-fold increase on average from baseline to peak compared to a 2-fold increase found in regular TSST studies [31]. Moreover, the demonstrated effect sizes of $\eta^2 = 0.37$ (Experiment 1), and $\eta^2 = 0.27$ (Experiment 2) go beyond the averaged effect sizes of $\eta^2 = 0.18$ found in a recent meta-analysis [31]. We therefore argue that the refined SECPT-G is a viable alternative to single session stress inductions because it reliably induces stress responses while allowing researchers to attain a high degree of standardization and reducing resource constraints.

Here, we focused on cortisol responses as previous research was ambiguous as to whether the SECPT-G could reliably affect the HPA axis. In comparison, SNS responses on cardiovascular are documented quite well, such as blood pressure and heart rate [12,15]. Thus, we decided to only assess sAA as an additional indicator for the SNS activity where a previous study did not report increased sAA levels following the SECPT-G [14]. Our results on sAA were mixed though: Although we found significant interactions, the pairwise comparisons between the stress conditions and the control conditions at the time directly following the manipulation were not significant. Thus, and due to the lack of power to find small effect sizes with the given sample sizes, the effects of the SECPT-G on sAA should be interpreted with care.

Despite this, our findings suggest that participants subjectively experienced more stress following the SECPT-G. Although this difference was not significant in Experiment 1, it could be argued that this may have resulted from assessing subjective stress with 60 items repeatedly on every measurement point. That is, the prolonged assessment of subjective stress might have exhausted participants and biased their responses. Indeed, when only assessing subjective stress with the VAS [18] Experiment 2 showed the expected effects: Both stress conditions were perceived as more stressful compared to the control condition directly after exposure to the stressor.

Importantly, the group-observation and individual-observation conditions did not differ in Experiment 2 regarding all outcome measures. These findings suggest that the personnel required to conduct the protocols can be reduced to one stress confederate (at least per 6 participants). Group-observation protocols may thus be used to decrease resource constraints and coordination effort. Thereby, the group-observation protocol also ensures a higher degree of standardization within an experiment.

Notably, the SECPT-G in a group-observation setting is also less resource-intensive regarding time and personnel, and more standardized in comparison to a group-administered TSST (TSST-G [11];: The TSST-G invites six participants to do a public mock job interview and an arithmetic task in which every participant has to present in front of the other participants, with two confederates being introduced as members of a selection committee (i.e., stress confederates). Here, participants on average show a 3-fold increase to the baseline following cortisol levels following the TSST-G [11]—being comparable to the effects shown with the SECPT-G. However, the TSST-G protocol takes 30 instead of 5 min to induce stress in the same number of participants and requires an additional stress confederate compared to the group-observation SECPT-G. To handle the situation and contain neutral demeanor, experimenters in the TSST-G also require substantial training. The SECPT-G thus offers an alternative to induce stress with fewer resource constraints, less required coordination and a higher degree of standardization even compared to other group-administered stressors.

4.1. Limitations and future research

Because SECPT protocols in single and group settings reliably affected the SNS (e.g. [1,10,12,15], we aimed to demonstrate that the SECPT-G can robustly activate the HPA axis. Consequently, we only assessed the SNS activity via sAA, but not via cardiovascular measures. Thereby, we did not need to bother participants with the application of the cardiovascular sensing equipment, enabling us to focus on cortisol and sAA as the main variables of interest.

In addition, it has to be noted that we did not investigate whether a group-observation condition also works when stressing 12 participants simultaneously, as was done in the protocol by Minkley et al. [12]. It can be argued that this would increase efficiency even further as more participants could complete the experiment at the same time. Here, we decided to test the group-observation SECPT-G by sampling six participants simultaneously because we presumed that this would reduce effort of coordinating participants and that rooms that can accommodate six participants (compared to twelve) are more readily available in most labs. Nonetheless, we believe that it is possible to implement the SECPT-G with twelve participants in a group-observation condition. Because we have no data for the group-observation with 12 participants and one stress confederate reasoning, we would recommend employing at least two stress confederates in such a scenario. This way, the same ratio of participants to stress confederate (6:1) can be attained and a similarly socially evaluative situation is created as in our Experiment 2. Nonetheless, future research should test whether the SECPT-G can also be successfully used with a different participant to stress confederate ratio.

5. Conclusion

In the present investigation, we present a refined template for the SECPT-G which reduces resource constraints (regarding personnel and time) compared to the individual SECPT, and can be conducted in a highly standardized manner. Our results show that this protocol is able to elicit robust stress responses in a group and in an individual-observation condition. To achieve this, it is important to incorporate socially evaluative and uncontrollable elements [8], such as a prominent positioning of cameras, and a lack of information about the duration of the stressor.

Author contributions

Benjamin Buttlar: Conceptualization, Methodology, Validation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Supervision; Helena Dieterle: Conceptualization, Methodology, Formal Analysis, Investigation, Writing – Review & Editing, Visualization; Regan Mandryk: Conceptualization, Methodology, Software, Validation, Resources, Writing – Review & Editing, Supervision, Project Administration.

Declarations of interest

None.

Funding

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

References

[1] L. Schwabe, L. Haddad, H. Schachinger, HPA axis activation by a socially evaluated cold-pressor test, Psychoneuroendocrinology 33 (2008) 890–895, https://doi.org/10.1016/j.psyneuen.2008.03.001.
[2] S.D. Preston, T.W. Buchanan, R.B. Stansfield, A. Bechara, Effects of anticipatory stress on decision making in a gambling task, Behav. Neurosci. 121 (2007) 257–263, https://doi.org/10.1037/0735-7044.121.2.257.
[3] A. Singh, J.S. Petrides, P.W. Gold, G.P. Chrousos, P.A. Deuster, Differential hypothalamic-pituitary-adrenal axis reactivity to psychological and physical stress, J. Clin. Endocrinol. Metab. 84 (1999) 1944–1948, https://doi.org/10.1210/jcem.84.6.5746.
[4] U.M. Nater, N. Rohleder, Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: current state of research, Psychoneuroendocrinology 34 (2009) 486–496, https://doi.org/10.1016/j.psyneuen.2009.01.014.
[5] C. Kirschbaum, D.H. Hellhammer, Salivary cortisol in psychobiological research: an overview, Neuropsychobiology 22 (1989) 150–169, https://doi.org/10.1159/000118611.
[6] E.A. Hines, G.E. Brown, The cold pressor test for measuring the reactibility of the blood pressure: data concerning 571 normal and hypertensive subjects, Am. Heart J. 11 (1936) 1–9, https://doi.org/10.1016/S0002-8703(36)90370-8.
[7] C. Kirschbaum, K.M. Pirke, D.H. Hellhammer, The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting, Neuropsychobiology 28 (1993) 76–81, https://doi.org/10.1159/000119004.
[8] S.S. Dickerson, M.E. Kemeny, Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research, Psychol. Bull. 130 (2004) 355–391, https://doi.org/10.1037/0033-2909.130.3.355.
[9] I. Labuschagne, C. Grace, P. Rendell, G. Terrett, M. Heinrichs, An introductory guide to conducting the trier social stress test, Neurosci. Biobehav. Rev. 107 (2019) 686–695, https://doi.org/10.1016/j.neubiorev.2019.09.032.
[10] L. Schwabe, H. Schächinger, Ten years of research with the socially evaluated cold pressor test: data from the past and guidelines for the future, Psychoneuroendocrinology 92 (2018) 155–161, https://doi.org/10.1016/j.psyneuen.2018.03.010.
[11] B. von Dawans, C. Kirschbaum, M. Heinrichs, The Trier Social Stress Test for Groups (TSST-G): a new research tool for controlled simultaneous social stress exposure in a group format, Psychoneuroendocrinology 36 (2011) 514–522, https://doi.org/10.1016/j.psyneuen.2010.08.004.
[12] N. Minkley, T.P. Schröder, O.T. Wolf, W.H. Kirchner, The socially evaluated cold-pressor test (SECPT) for groups: effects of repeated administration of a combined physiological and psychological stressor, Psychoneuroendocrinology 45 (2014) 119–127, https://doi.org/10.1016/j.psyneuen.2014.03.022.
[13] J.A. Häusser, M. Kattenstroth, R. van Dick, A. Mojzisch, "We" are not stressed: social identity in groups buffers neuroendocrine stress reactions, J. Exp. Soc. Psychol. 48 (2012) 973–977, https://doi.org/10.1016/j.jesp.2012.02.020.
[14] L. Becker, U. Schade, N. Rohleder, Evaluation of the socially evaluated cold-pressor group test (SECPT-G) in the general population, PeerJ 7 (2019) e7521, https://doi.org/10.7717/peerj.7521.
[15] B. von Dawans, A. Trueg, C. Kirschbaum, U. Fischbacher, M. Heinrichs, Acute social and physical stress interact to influence social behavior, The role of social anxiety. PLOS ONE 13 (2018), e0204665, https://doi.org/10.1371/journal.pone.0204665.
[16] F. Faul, E. Erdfelder, A.-G. Lang, A. Buchner, G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences, Behav. Res. Methods 39 (2007) 175–191, https://doi.org/10.3758/BF03193146.
[17] B.M. Kudielka, D.H. Hellhammer, S. Wüst, Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge, Psychoneuroendocrinology 34 (2009) 2–18, https://doi.org/10.1016/j.psyneuen.2008.10.004.
[18] P. Zimmer, B. Buttlar, G. Halbeisen, E. Walther, G. Domes, Virtually stressed? A refined virtual reality adaptation of the Trier Social Stress Test (TSST) induces robust endocrine responses, Psychoneuroendocrinology 101 (2019) 186–192, https://doi.org/10.1016/j.psyneuen.2018.11.010.
[19] C. Harmon-Jones, B. Bastian, E. Harmon-Jones, The Discrete emotions questionnaire: a new tool for measuring state self-reported emotions, PLoS One 25 (2016).
[20] C. Spielberger, R. Gorsuch, R. Lushene, P. Vagg, G. Jacobs, Manual for the State-Trait Anxiety Inventory (Form Y1 – Y2), Consulting Psychologists Press, Palo Alto, CA, 1983.
[21] M.F. Scheier, C.S. Carver, The self-consciousness scale: a revised version for use with general Populations1, J. Appl. Soc. Psychol. 15 (1985) 687–699, https://doi.org/10.1111/j.1559-1816.1985.tb02268.x.
[22] G.J. Boyle, Self-description questionnaire II: a review, Test Crit 10 (1994) 632–643.
[23] E. Deci, R. Ryan, The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. Psychol, Inq 11 (2000) 227–268, https://doi.org/10.1207/S15327965PLI1104_01.
[24] C.S. Carver, You want to measure coping but your protocol's too long: consider the brief COPE, Int. J. Behav. Med. 4 (1997) 92–100, https://doi.org/10.1207/s15327558ijbm0401_6.
[25] R.M. Ryan, Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory, J. Pers. Soc. Psychol. 43 (1982) 450–461, https://doi.org/10.1037/0022-3514.43.3.450.
[26] S.A. Jackson, H.W. Marsh, Development and validation of a scale to measure optimal experience: the Flow state scale. J. Sport exerc, Psychol. 18 (1996) 17–35, https://doi.org/10.1123/jsep.18.1.17.
[27] S. Sonnentag, C. Fritz, The Recovery experience questionnaire: development and validation of a measure for assessing recuperation and unwinding from work, J. Occup. Health Psychol. 12 (2007) 204–221, https://doi.org/10.1037/1076-8998.12.3.204.
[28] G. Chen, S. Gully, D. Eden, Validation of a new general self-efficacy scale. Organ, Res. Methods - ORGAN RES METHODS 4 (2001), https://doi.org/10.1177/109442810141004.
[29] E.R. Girden, ANOVA: Repeated Measures, ANOVA: Repeated Measures, Sage Publications, Inc, Thousand Oaks, CA, US, 1992.
[30] J.C. Pruessner, C. Kirschbaum, G. Meinlschmid, D.H. Hellhammer, Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change, Psychoneuroendocrinology 28 (2003) 916–931, https://doi.org/10.1016/s0306-4530(02)00108-7.
[31] W.K. Goodman, J. Janson, J.M. Wolf, Meta-analytical assessment of the effects of protocol variations on cortisol responses to the Trier Social Stress Test, Psychoneuroendocrinology 80 (2017) 26–35, https://doi.org/10.1016/j.psyneuen.2017.02.030.