Physiology & Behavior 242 (2021) 113618

Contents lists available at ScienceDirect

Physiology & Behavior

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

Habituation of salivary cortisol and cardiovascular reactivity to a repeated real-life and virtual reality Trier Social Stress Test

Oswald D. Kothgassner [a,b,c], Andreas Goreis [b,d], Lisa M. Glenk [e], Johanna Xenia Kafka [a,c,f], Bettina Pfeffer [a,b,c], Leon Beutl [g], Ilse Kryspin-Exner [b], Helmut Hlavacs [g], Rupert Palme [8], Anna Felnhofer [b,c,f,*]

[a] Department of Child and Adolescent Psychiatry, Medical University of Vienna, Vienna, Austria
[b] Department of Clinical and Health Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria
[c] Comprehensive Center of Pediatrics, Medical University of Vienna, Vienna, Austria
[d] Outpatient Unit for Research, Teaching and Practice, Faculty of Psychology, University of Vienna, Vienna, Austria
[e] Comparative Medicine, the interuniversity Messerli Research Institute, University of Veterinary Medicine Vienna, Medical University Vienna, University of Vienna, Austria
[f] Department of Pediatrics and Adolescent Medicine, Medical University of Vienna, Vienna, Austria
[g] Working Group Entertainment Computing, University of Vienna, Vienna, Austria
[8] Department of Biomedical Sciences, University of Veterinary Medicine, Vienna, Austria

ARTICLE INFO

Keywords:
Virtual reality
TSST
Salivary cortisol
Heart rate
Social Presence
Psychosocial Stressor

ABSTRACT

Background: Although the Trier Social Stress Test (TSST) constitutes a valid paradigm for social stress induction, less is known about the effects of a virtual reality (VR) TSST on short- and long-term hypothalamic-pituitary-adrenal (HPA) axis and sympathetic-adreno-medullar (SAM) axis responses. Hence, this study set out to evaluate reactivity and habituation of self-reported stress and HPA and SAM reactivity in a real TSST and VR-TSST when compared to a placebo TSST.
Method: Sixty-eight healthy young adults (50% female) were randomly assigned to either a real TSST, a VR-TSST, or a placebo TSST, all of which were conducted three times (one day and one week post initial exposure). Social presence, self-reported stress, salivary cortisol, heart rate (HR), and heart rate variability (HRV) were analyzed using ANOVAs and multilevel models.
Findings: On the first exposure, both the real and VR-TSST showed significantly stronger cortisol and cardiovascular responses than the placebo. On the second visit, the cortisol response was still significantly high—and the HRV response low—for the real and VR-TSST. The third visit resulted in HR, HRV, and cortisol responses comparable to the placebo group. Furthermore, the real TSST induced more self-reported stress than the placebo on all three visits, the VR-TSST only on the first two visits. Social presence was stable across conditions and had no association with stress markers.
Conclusion: These findings imply that the replicability of stress exposures at shorter intervals seems problematic for the traditional TSST, and for the VR-TSST.

1. Introduction

Over the past decades, the Trier Social Stress Test (TSST; [18]) has established itself as a consistent, reliable, and valid protocol for evoking psychosocial stress. The TSST consists of a mock job interview and an arithmetic task, and, as such it is one of the most widely used and extensively validated laboratory stress paradigms. Two factors are regarded as crucial in experimentally manipulating psychosocial stress: the social evaluative context and the uncontrollability of the situation [3]. Overall, the TSST has repeatedly been shown to evoke robust hypothalamic-pituitary-adrenal (HPA) axis and sympathetic-adreno-medullar (SAM) axis responses in most participants [10]. This effect has been found to be robust and stable over different adaptations of the protocol [8].

Although the original paradigm is intended for face-to-face

* **Corresponding author** at: Department of Pediatrics and Adolescent Medicine, Medical University of Vienna, Waehringer Guertel 18-20, Postfach 27, 1090 Vienna, Austria.
E-mail address: anna.felnhofer@meduniwien.ac.at (A. Felnhofer).

https://doi.org/10.1016/j.physbeh.2021.113618
Received 6 August 2021; Received in revised form 29 September 2021; Accepted 2 October 2021
Available online 5 October 2021

interactions, past research has also established its effectiveness in technologically mediated settings, such as immersive technologies (e.g., [22, 26]) or online contexts (TSST-OL; [9]). Both forms of applications are attractive in terms of simplification and standardization of the TSST protocol [11] and in terms of economically applying advanced measurements (e.g., eye-tracking; [34]). The immersive nature of virtual reality (VR) technologies may be significant as it allows social stressors to be presented in a standardized and ecologically valid manner. While only a few studies have, to date, directly compared real-world and VR based TSST protocols, preliminary findings (e.g., [19]) suggest that real and virtual stressors are comparable in their ability to induce not only high subjective stress levels but also to provoke considerable physiological stress reactivity on both HPA and SAM axes.

Yet, extant findings are conflicting: While Kelly et al. [16] and Shiban et al. [31] imply that a VR-TSST evokes SAM and self-reported stress responses, which are similar to a real-life TSST, they also found the VR-TSST to be associated with lower endocrine stress levels (cortisol) than the real TSST. Contrary to this, Zimmer et al. [35] demonstrated that stress-induced increases of free salivary cortisol and alpha-amylase were equal in both stress groups (VR and real) and significantly higher than in the control group [35]. A recent meta-analysis [11] adds further sustenance to the latter finding in that it shows that there are no significant differences in HPA, SAM, and self-reported reactivity between the real TSST and VR-TSST. This heterogeneous study situation could be explained by differences in social presence [1]. Experiencing the social presence of others is a predictor of anxiety and stress, particularly in socially anxious individuals [5]. As such, it could impact stress responses in technologically mediated TSSTs.

Additionally, there is a growing body of research on VR-TSSTs and their effect on immediate stress responses (see [11]), but the effect of repeatedly applying the VR-TSST on stress reactivity has so far been understudied. Indeed, the human ability to habituate to repeated exposures to the same stress stimulus is a crucial precondition for reducing allostatic load – a stage in which physiological systems are no longer able to adapt to stress [25]. This means that failing to habituate to stress exposure may lead to a higher vulnerability for mental disorders and somatic disease [15]. It is critical to assess these habituation effects in a standardized manner, and particularly the replicability of stress habituation for novel, technology-based methods such as the VR-TSST still needs to be determined. Furthermore, knowledge about the habituation to social stress caused by virtual entities would allow better predictions on the role of virtual social stressors and if these stressors affect psychobiological systems in a comparable way as real-life stressors do. In traditional TSSTs, HPA reactivity is known to gradually decrease after repeated stress provocations – most notably during the second visit –, while autonomic and self-reported stress measures remain relatively stable [30]. To our knowledge, there is only one pilot study by Jönsson et al. [14] who used immersive CAVE technology in a small number of participants ($n = 10$) and as such provided the first indication of this habituation effect also with immersive technologies. However, a direct standardized comparison to a real TSST is still outstanding.

Given this paucity of studies, the present research aims to investigate the psychobiological reactivity and habituation of the HPA and SAM axes as well as self-reported stress levels in repeated exposures to a real TSST versus a VR-TSST. We also introduce a placebo TSST according to Het et al. [12] to control for the effects of technology and social stress induction. We hypothesize that we will find similar HPA and SAM reactivity and habituation to the real TSST and the VR-TSST but significant differences to the placebo TSST.

2. Material and methods

2.1. Participants

Participants were recruited through advertisements at the University of Vienna and on social media platforms. Our sample consisted of $N = 68$, of which $n = 34$ were female ($M_{age} = 3.82$, $SD_{age} = 2.74$) and $n = 34$ were male ($M_{age} = 24.44$, $SD_{age} = 2.70$); males and females were balanced across groups. The mean body mass index (BMI) was 22.46 ($SD = 2.33$). The absence of mental disorders in our sample was determined using the German Version of the Structured Interview for DMS-IV (SCID; [7]), a semi-structured interview which was conducted by trained personnel. None of the participants reported having used drugs for the last 12 months or medication for 2 weeks prior to their participation in the study. As participants came to the lab on three separate days (see "Procedure" for details), they were instructed to refrain from alcohol, caffeine, or exercise 24 h prior to the day of each assessment. A total of 18 participants (26% of the sample) reported to have smoked cigarettes in the past four weeks (mean number of cigarettes: 32, $SD = 31$); however, all of these participants were abstinent from nicotine in the 72 h preceding all three visits, and habitual smoking did not have an effect on analyses of salivary cortisol at all three visits ($ps > 0.477$). The groups did not differ regarding their technical expertise ($p = .936$) measured on a single-item questionnaire on a 5-point Likert-scale. All female participants participated during their luteal phase in the menstrual cycle for a better comparison with male participants; furthermore, we excluded females using oral contraceptives from the study [17].

2.2. Procedure

The current study was conducted according to the Declaration of Helsinki, and thus a comprehensive informed consent was obtained from all participants prior to their voluntary participation. All participants were allowed to withdraw from the experiment at any time. All participants volunteered in exchange for course credit. The experimental procedure took place between 15:00 and 17:00 in temperature-controlled (23 °C) rooms. Participants were invited for the exact timeframe for each session to control for the circadian rhythm [21]. Participants were randomly assigned to either A) a real TSST condition ($n = 22$), B) a VR-TSST condition ($n = 23$), or C) a placebo TSST condition ($n = 23$). Both the real TSST and the VR-TSST were parallelized regarding the setting and the elements of the psychosocial stressor. Participants underwent their respective condition and were subsequently invited to two more appointments (1 day after initial TSST and 1 week after initial TSST; with same timing of measurement to control for daily rhythm of cortisol levels in saliva) to undergo the same condition again twice (referred from hereon as visits 1–3). Following other studies, the job description was slightly changed during the second and third visits [14, 30].

As depicted in **Fig. 1**, the procedure for each visit consisted of a preparation period upon the arrival at the lab (30 min), the TSST (20 min), a post-TSST period (10 min) as well as a resting (40 min) and debriefing period (10 min). During the preparation and the post-TSST period, psychometric assessments were conducted, and physiological measures were applied. The TSST was used to induce psychosocial stress. Participants were standing in an upright position during baseline recording, TSST, and post-TSST period. The resting and debriefing periods were conducted in a sitting position.

2.3. Instruments

2.3.1. Fear of negative evaluation

Fear of negative evaluation was assessed with the Fear of Negative Evaluation Scale (FNE, [23]), using 12 items (e.g., "I am afraid that others will not approve of me"). Cronbach's Alpha was 0.79 in our sample. Responses were measured on a 5-point Likert scale (strongly agree—strongly disagree).

2.3.2. Subjective stress

Subjective stress was rated via single-item visual analog scales (VAS) at four time points (see **Fig. 1**). Participants were asked to rate "how stressed" they felt at a given moment. Answers were marked on a line of

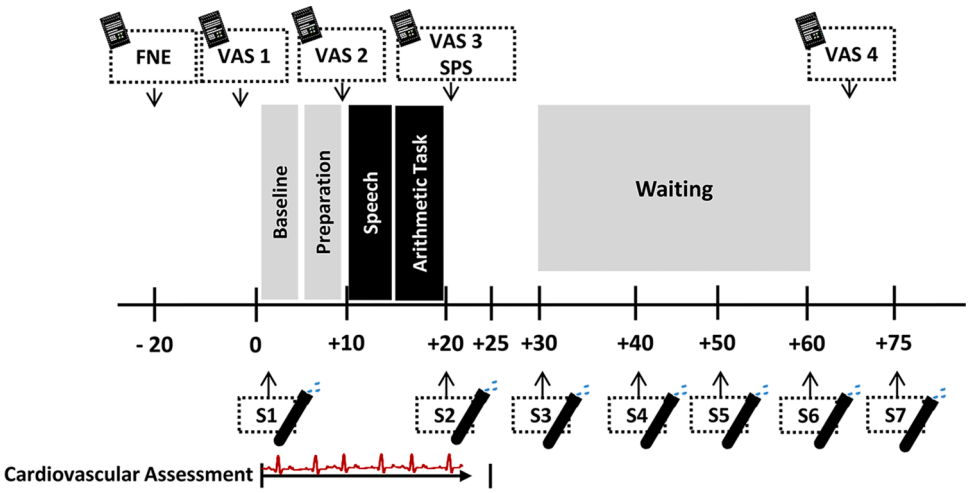

Fig. 1. Experimental design of each of the three TSST-visits. *Note:* FNE = Fear of negative evaluation, VAS = Visual Analogue Scales, SPS = Social Presence Scale, *S* = Salivary cortisol.

100 mm with endpoints marked as 0 = not at all and 100 = extremely.

2.3.3. Social presence

The five-item Social Presence Survey (SPS; [1]) was used to assess social presence experienced during the VR-TSST at the end of all three visits for a total of three times. Example items include 'The persons appeared to be sentient, conscious, and alive to me'. Responses were measured on a 7-point Likert scale (strongly agree—strongly disagree).

2.3.4. TSST procedures

The real-life TSST. This TSST was based on the widely used procedure by Kirschbaum et al. [18]. In a 5 min preparation period, all participants were instructed to prepare for a job application and introduce themselves to a committee of two interviewers (one female and one male). Following this period, the participants completed a 5 min mock job interview and a 5 min arithmetic task (serial subtraction in steps of 13 starting at 1022) in front of the interviewers.

The VR-TSST. As in the real-life TSST, the participants were instructed to prepare for a job application and underwent a TSST procedure in a virtual environment. The VR-TSST (see Fig. 2) was developed using an open-source engine (Ogre3D) for the real-time rendering of the scenes. GIMP and *C++* were used for source code, textures, and graphical surfaces. Participants were donned a Head-Mounted Display (HMD; Sony HMZ-T1, Sony, Japan) with an external head tracking device (Track IR5, Natural Point, US). We carefully parallelized both conditions across all visits.

The Placebo-TSST. Further, we administered a non-stressful placebo version of the TSST according to Het et al. [12] in a virtual reality condition (no social evaluation) to control for possible influences of the VR apparatus. Therefore, participants had to talk about a self-chosen topic without the presence of social entities or evaluative cues in the virtual room. This procedure also allowed us to investigate the influence of social stimuli on stress outcomes.

2.3.5. Electrophysiological measures

Heart rate (HR; in bpm) was measured as a marker of stress-related sympathetic activity, and Heart Rate Variability (HRV) as the Root Mean Square of Successive Differences (RMSSD; in ms) by continuous recording for subsequent 60 s intervals from the baseline measure 5 min before preparation for TSST until 5 min after the active TSST phase. We

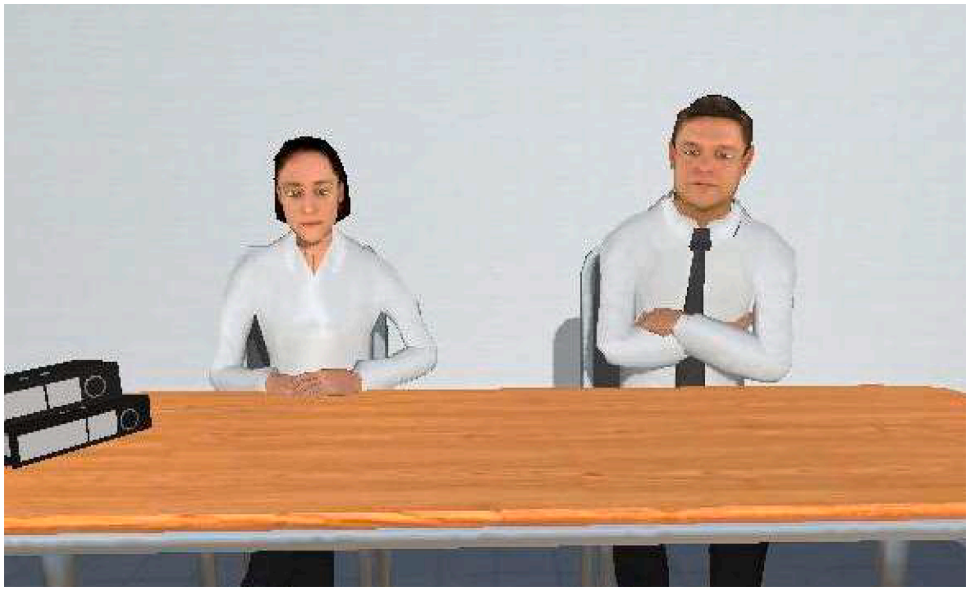

Fig. 2. VR-TSST with a female and a male jury member.

monitored cardiovascular data via a portable M-EXG (Schuhfried, Mödling, Austria) with a sample rate of 1000 Hz. One-way electrodes (Medica RedDot Electrodes, Perchtoldsdorf, Austria) were located on the seventh intercostal space on the right and left sides of the body to measure cardiovascular activity. Data were computed using the KUBIOS HRV software kit (Biosignal Analysis and Medical Imaging Group, Finland). High HR values indicate more bpm and high physiological arousal, whereas low HR values reflect less bpm and lower physiological arousal. Conversely, low RMSSD values indicate lower parasympathetic activity (therefore, higher arousal), and high RMSSD values reflect more parasympathetic activity, indicating lower physiological arousal.

2.3.6. Salivary cortisol

We used commercial cotton swabs (Salivette®, Sarstedt, Wiener Neudorf, Austria) without any saliva-stimulating additives for saliva collection. Participants were thoroughly instructed on how to collect their saliva on their own. Thus, participants put the swab into their cheek pouch at designated time points and let the swab saturate with saliva for approximately 60 to 80 s. Afterwards, the swabs were replaced into the device container and immediately frozen at -20 °C. For cortisol analysis, samples were thawed on ice and centrifuged at room temperature (3000 × g; 15 min). Samples were assayed (after a 1 + 9 dilution with assay buffer) in duplicates with a cortisol enzyme immunoassay [27]. Please see Kothgassner et al. [20] for further details about the analysis.

2.3.7. Statistical analysis

Data were analyzed using R 4.0.2 [28], considering a significance level of $p < .05$. To test for differences in the psychological and physiological stress responses, we computed multilevel models (MLMs, [6]) using the lme4 package [2]. We used MLMs, as they are recommended over other means of analysis (e.g., repeated measures ANOVA) for the analyses of biological repeated measures data, given that MLMs are not vulnerable to missing values and do not require homogeneous covariance structures [13]. In these MLMs, individual time points were nested into visits. As our goal was to conceptualize and compare stress responses between the three conditions, we used models that included individual time points, condition (dummy coded as 0 = placebo TSST, 1 = real TSST, 2 = VR-TSST, where the placebo TSST was the reference category), and the condition × time interaction for each visit separately. For the salivary cortisol models, we also included time^2 (i.e., curvilinear deceleration) and the interaction of condition × time^2. Simple effects analyses were used to break down significant interactions. Participant id was included as a random effect (random intercept), and the age of participants was included as a covariate in all models.

Salivary cortisol (Shapiro-Wilk $p < .001$) was not normally distributed; to reduce skewness, we, therefore, log-transformed it prior to analyses. In sum, 8% of all data was missing. We did not impute missing values or exclude these participants as MLMs can be applied with incomplete cases [33]. We planned our sample size to find a large effect size according to former studies (e.g., [11]; Zimmer et al., 2019) in stress responses during the real and VR-TSST compared to the placebo TSST. Using power analysis simulations with the package powerlmm [24], we found that our study was sufficiently powered to estimate 2-way interactions (i.e., condition x time) with a large effect size ($\beta = 0.50$) assuming 7 level-1 observations for salivary cortisol, 5 level-1 observations for HR, and 4 level-1 observations for subjective stress, all nested within 23 level-2 observations (participants) per condition. In detail, the power (assuming $\alpha = 0.05$) was 93% for salivary cortisol, 99% for cardiovascular reactivity, and 84% for subjective stress.

3. Results

3.1. Fear of negative evaluation and baseline comparisons of physiological and psychological outcomes

Group comparisons revealed that participants in the three conditions did not differ regarding their FNE scores (real TSST: $M = 31.23$, $SD = 8.47$; VR-TSST: $M = 34.78$, $SD = 7.85$; placebo TSST: $M = 34.00$, $SD = 2.59$; $F = 1.69$, $p = .193$). Furthermore, baseline comparisons of the first measurement points of the three visits showed that the three groups did not differ significantly in their baseline levels of salivary cortisol (visit 1: $F = 0.52$, $p = .596$; visit 2: $F = 0.93$, $p = .400$; visit 3: $F = 0.30$, $p = .744$), HR ($F = 0.02$, $p = .977$; $F = 0.33$, $p = .968$; and $F = 0.04$, $p = .996$), HRV ($F = 0.61$, $p = .548$; $F = 0.78$, $p = .464$; and $F = 1.33$, $p = .276$), and self-reported stress ($F = 0.11$, $p = .897$; $F = 0.83$, $p = .440$; and $F = 1.15$, $p = .323$). Correlations among stress responses across visits are depicted in the supplementary material.

3.2. Salivary cortisol

At visit 1, we found that the increase of salivary cortisol (time predictor) was significantly steeper in the real TSST ($b = 0.037$, $p < .001$; increase from baseline: 95.2%) and VR-TSST ($b = 0.026$, $p = .014$; increase from baseline: 73.3%) when compared to the placebo TSST (which was coded as the reference category). The deceleration (time^2 of cortisol was more pronounced in the real TSST ($b = -0.0005$, $p = .002$), followed by the VR-TSST ($b = -0.004$, $p = .021$). No difference between the real and virtual conditions was found regarding the increase ($p = .302$) and deceleration ($p = .367$) of cortisol at visit 1. At visit 2, the linear increase of cortisol was again significantly steeper in the real TSST ($b = 0.030$, $p = .002$; increase from baseline: 21.6%), and the VR-TSST ($b = 0.029$, $p = .002$; increase from baseline: 9.4%) – both when compared to the placebo condition. The deceleration of cortisol was significant in the real TSST ($b = -0.0004$, $p = .006$) and the VR-TSST ($b = -0.0005$, $p = .001$). Similar to visit 1, no differences between the real and virtual conditions were found regarding the increase ($p = .916$) and deceleration ($p = .687$) at visit 2. At visit 3, neither the real (time: $p = .926$, time^2: $p = .826$) nor the VR-TSST (time: $p = .941$, time^2: $p = .985$) condition showed a linear increase or a downward deceleration of cortisol different from the placebo TSST. In summary, the real and the VR-TSST induced a cortisol response at both the first and second visit—with the real TSST showing a stronger linear increase in cortisol values. The third iteration did not induce any cortisol response in the real or VR-TSST. All results are depicted in **Fig. 3**.

3.3. Heart rate

HR trajectories are shown in **Fig. 4A**. In our analysis of HR, we found a significant interaction of condition x time at visit 1 ($F = 4.89$, $p < .001$). Simple effects analyses indicated that the real ($t = 2.63$, $p < .027$, increase from baseline: 22.11 bpm) and VR-TSST ($t = 2.41$, $p = .047$, increase from baseline: 17.97 bpm) conditions had a higher HR during the TSST (i.e., 5 min after the start of the TSST) as compared to the placebo TSST (the increase did not differ between real and virtual; $p = .946$). At visit 2, an interaction of time × condition was also found ($F = 4.91$, $p = <0.001$) and simple effect analyses indicated that only the real ($t = 2.48$, $p = .041$, increase from baseline: 17.20 bpm) but not the VR-TSST ($t = 2.29$, $p = .064$, however: increase from baseline: 14.56 bpm) had a significantly higher HR during the TSST than then placebo TSST. On the third visit, a significant condition x time interaction was found as well ($F = 2.24$, $p = .027$), but there were no differences regarding HR between the real (increase from baseline: 11.45 bpm), virtual (increase

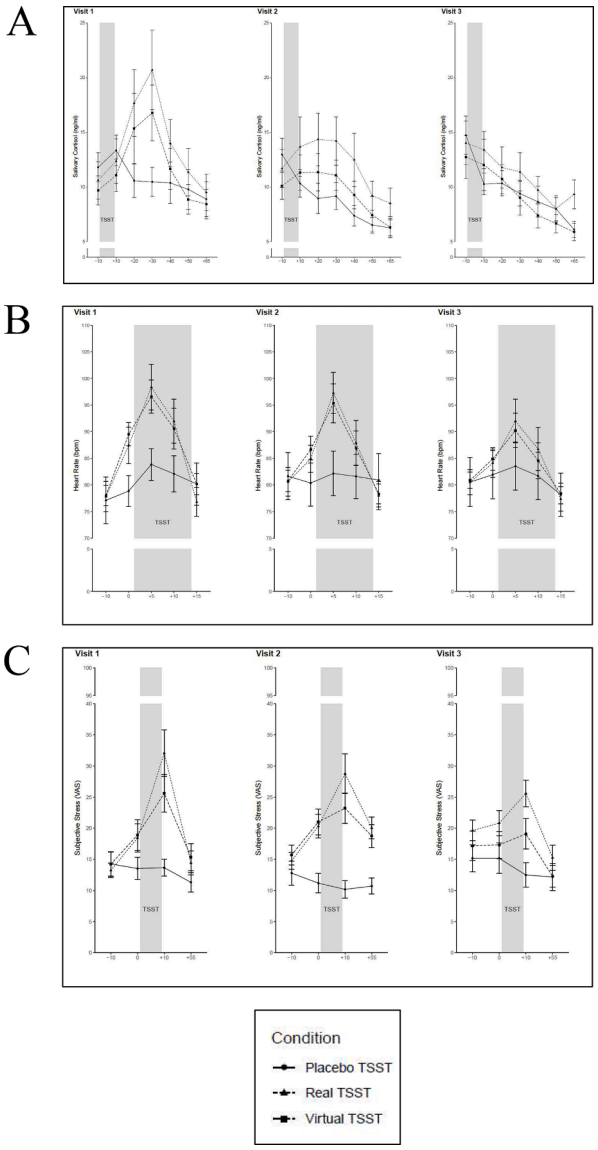

Fig. 3. Salivary Cortisol of three conditions across the three visits (Mean ± SEM).

from baseline: 9.34 bpm), and placebo TSST during stress induction (*p*s > 0.605). In summary, in the first iteration of TSSTs, both the real and the VR-TSST induced a significant increase in HR. Only the real TSST induced a significant HR response during the second visit, whereas the VR-TSST condition exhibited an increase from baseline to peak of 14.56, which, however, was not significant. During the third time, no significant increase in HR as a response to the TSSTs was found.

3.4. Heart rate variability

HRV trajectories are shown in Fig. 4**B**. For the parameter RMSSD, we found a significant interaction of condition × time at visit 1 ($F = 5.02$, $p < .005$). Simple effects analyses showed that the real TSST ($t = 3.16$, $p = .007$, decrease from baseline: −20.98 ms) and the VR-TSST ($t = 3.05$, $p = .009$, decrease from baseline: −13.19 ms) conditions had a lower RMSSD during the stress exposure as compared to the placebo TSST. The decrease of RMSSD did not differ between the real and virtual TSST at visit 1 ($p = .956$). At visit 2, we also found an interaction between condition × time ($F = 2.06$, $p = .042$) with simple effects, again, indicating that the real ($t = 3.01$, $p = .010$, decrease from baseline: -14.38 ms) and the VR-TSST ($t = 2.68$, $p = .024$, decrease from baseline: -8.53 ms) had a more attenuated HRV during the TSST than the placebo TSST (this decrease did not differ between the real and VR-TSST; $p = .865$). On the third iteration of the TSST, no significant interaction between condition × time was found ($F = 0.63$, $p = .628$). In summary, at the first and second TSSTs, both the real and the VR-TSST induced a significant decrease in RMSSD. At the third iteration, no significant decrease in RMSSD as a response to the TSSTs was found in either condition.

3.5. Subjective stress

A plot of subjective stress ratings across the three visits is depicted in Fig. 5. On the first visit, time and condition interacted ($F = 6.68$, $p < .001$). Simple effects analyses showed that on the first visit, a significant increase regarding subjective stress was found 10 min after the TSST in the real condition ($t = 6.03$, $p < .001$, increase from baseline: 18.86; on a scale from 0 to 100) and the virtual condition ($t = 3.98$, $p < 001$, increase from baseline: 11.39) when compared to the placebo TSST. Real and VR-TSSTs did not differ significantly regarding the magnitude of elicited subjective stress in the first visit ($p = .084$). Similarly, a time × condition interaction was found at the second visit ($F = 7.86$, $p < .001$), with higher subjective stress in the real ($t = 7.06$, $p < .001$, increase from

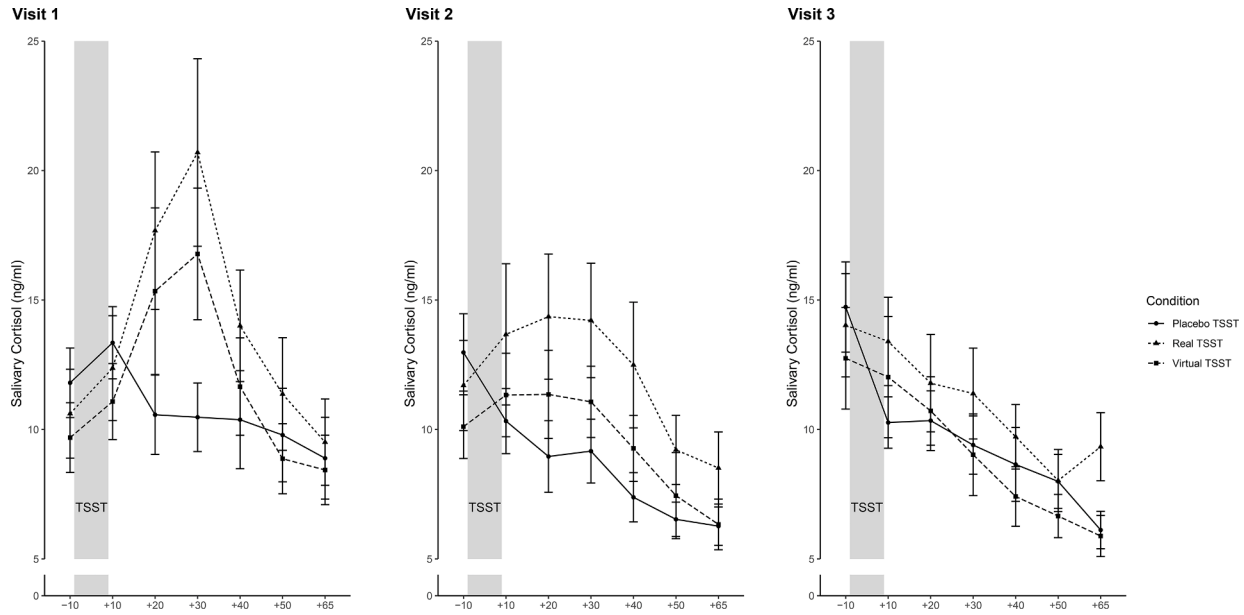

Fig. 3. (*continued*).

baseline: 14.00) and VR-TSST ($t = 5.03$, $p < .001$, increase from baseline: 7.52) when compared to the placebo TSST. Again, no differences regarding subjective stress were found between real and VR-TSSTs at visit 2 ($p = .080$). On the third iteration of the TSST, a significant time × condition interaction ($F = 2.67$, $p = .016$) was found, yet only the real TSST ($t = 4.26$, $p < .001$, increase from baseline: 5.96), not the VR-TSST ($t = 2.16$, $p = .082$, increase from baseline: 1.90) elicited more subjective stress than the placebo TSST. Summing up, the real TSST induced more subjective stress than the placebo TSST at all three visits, the VR-TSST only on the first two visits.

3.6. Social presence in the VR condition

Results from the SPS indicated that the perceived social presence did not differ over time in the VR TSST conditions as the main effect of time was not significant ($F = 1.390$, $p = .260$: visit 1: $M = 26.34$, $SD = 5.53$; visit 2: $M = 26.18$, $SD = 5.17$, visit 3: $M = 28.00$, $SD = 5.67$). SPS scores furthermore did not predict cortisol increase (time × SPS; visit 1: $p = .627$, visit 2: $p = .590$, visit 3: $p = .338$) or deceleration (time^2 × SPS; visit 1: $p = .714$, visit 2: $p = .419$, visit 3: $p = .578$), HR response (i.e., correlation with delta values, visit 1: $p = .902$, visit 2: $p = .772$, visit 3: $p = .540$), HRV response (visit 1: $p = .773$, visit 2: $p = .807$, visit 3: $p = .918$), or increase in subjective stress (visit 1: $p = .194$, visit 2: $p = .546$, visit 3: $p = .641$) at any of the three visits in the VR condition.

4. Discussion

This study aimed to investigate the reactivity and habituation of the HPA and SAM axis, and self-reported stress in response to a repeated real or virtual psychosocial stressor. Hence, we compared a traditional TSST with a VR-TSST and a placebo TSST regarding their effects on salivary cortisol, HR, HRV, and self-reported stress levels.

4.1. Initial stress response

Our results show a comparable increase in salivary cortisol regarding the real TSST and the VR-TSST on the first visit; both were significantly different from the placebo group. There was a marked increase (95.2% vs. baseline) during the first stress exposure in the real TSST and a 73% increase vs. baseline in the VR-TSST. These results are in line with previous studies [35] as well as with the meta-analysis by Helminen et al. [11] regarding single stress exposures. Similar to the HPA axis, the real and VR-TSST also seem to induce comparable SAM axis reactivity on the first exposure. While in the study by Zimmer et al. [35], SAM axis responses were slightly less pronounced in the VR than in the real TSST, our results regarding SAM axis reactivity – together with significant differences in self-reported stress – are consistent with Kelly et al. [16] and Shiban et al. [31]. A possible explanation of these findings might be that participants in our study reported relatively strong social presence experiences (mean of 5.2 to 5.6 on a 7-point Likert-Scale) in all three visits. Therefore, we conclude that participants in our study considered the two virtual agents as a convincing social-evaluative threat, which explains the increased HPA reactivity to the VR-TSST. A recent study [4] underlines that believability affects the effectiveness of VR-TSST, as it reports that participants who believed the VR-TSST setting showed increased HR levels. Another explanation is that in our sample most participants reported to have a considerable technical expertise (mean of 4.23 to 4.30 across groups on a 5-point Likert-Scale); this could particularly help counteract the increase in stress reactivity due to the novelty of the virtual environment as discussed by Zimmer et al. (2019).

4.2. Habituation of stress response

In our study, significant effects on salivary cortisol secretion were observed on the first two visits for both the real TSST and VR-TSST compared to placebo. However, on the second visit, the increase was much smaller for the real TSST (increase: 21%) and the VR-TSST (increase: 9.4%), indicating that some form of habituation to the stressor had taken place. Significant HR increases were also observed on the first two visits, although only for the traditional TSST. Despite an increase of HR in the VR-TSST on the second visit, there was no significant difference in HR when compared to the placebo group, even if the increase showed a tendency towards significance ($p = .064$). Furthermore, significant decreases in HRV were found at the first two visits with a comparable magnitude in both the real and the VR-TSST. On the third visit – one week after the first exposure – there were no effects in stress reactivity, neither in salivary cortisol, HR, or HRV. Surprisingly, the trajectories even decreased in all groups, like in the placebo TSST. Regarding subjective stress, participants reported higher levels in the real TSST than in the placebo TSST at all three time points; the same applied to the VR-TSST only on the first two visits.

These results are particularly important as, to our knowledge, there

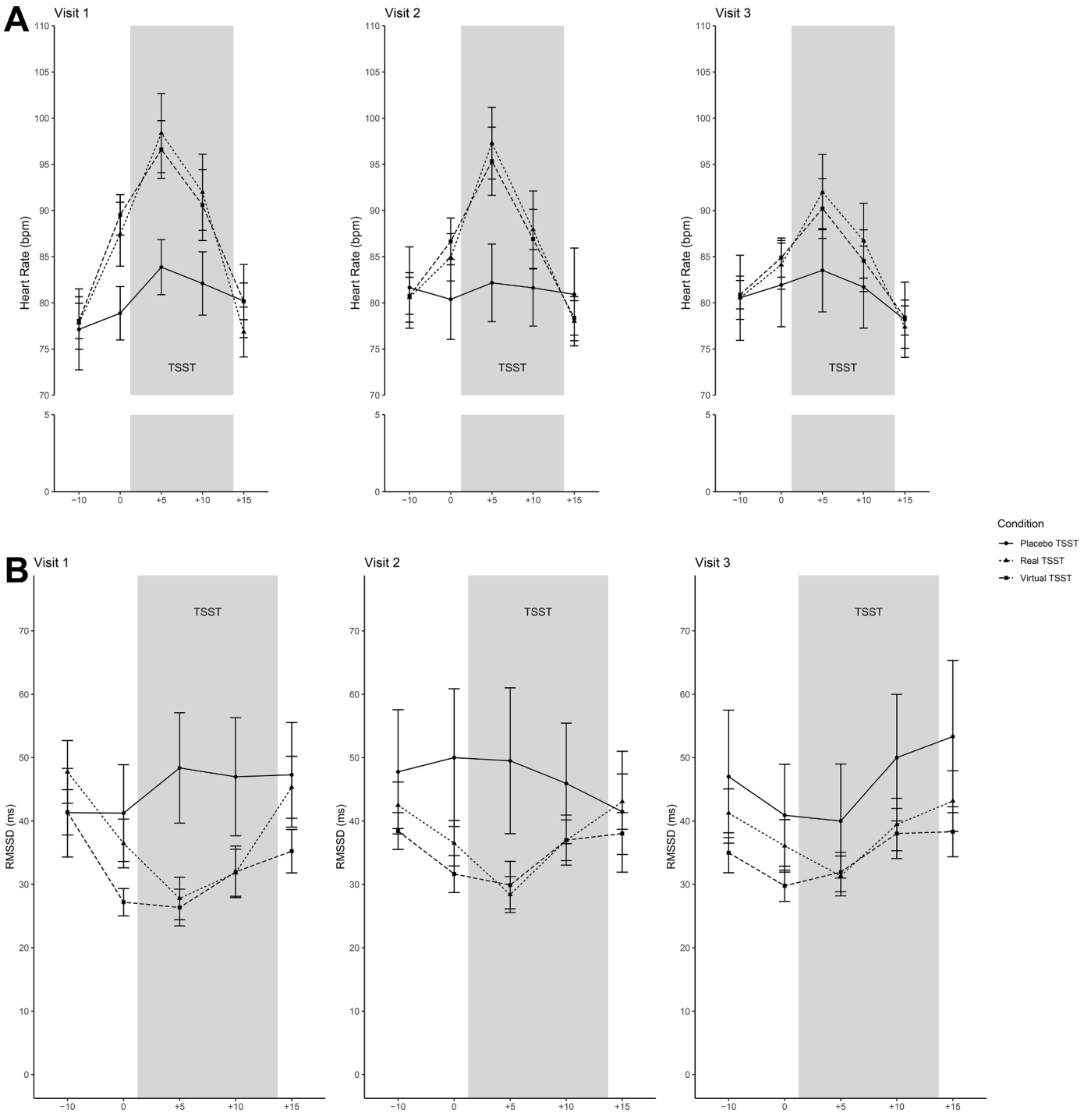

Fig. 4. Heart Rate (A) and Heart Rate Variability (B) of three conditions across the three visits (Mean ± SEM).

is only one uncontrolled small pilot study so far on the habituation to a VR-TSST [14]. It suggests that, compared to a control group, habituation of the effect is likely with multiple presentations of a TSST, particularly with the VR-TSST. Their results showed that the VR-TSST no longer differed from placebo on a second visit, at least in the autonomic stress response. At the third repetition of our experiment, there was no reactivity for neither the real nor the VR-TSST. This is especially interesting because multiple exposures under the same conditions and in a familiar setting may corrupt the effect of both real and VR-TSST. While Schommer et al. [30] allowed four weeks to elapse between the three runs and found habituation effects, we implemented a stress exposure with shorter time intervals, but still with comparable results.

4.3. Role of technology and social presence

The differences across studies may be attributable to the immersiveness of used technologies [32]. For instance, Jönsson et al. [14] used a CAVE-system that is less immersive in contrast to a head-mounted-display (HMD). Past research has provided sustenance to this: Montero-López et al. (2016) exposed their participants to a VR-TSST either via goggles or on a large screen and found the sympathetic stress reactivity (i.e., skin conductance levels) to be larger for VR-glasses. Salivary cortisol, however, was the same across both conditions. While a direct comparison between our and Jönsson and colleagues' (2010) data is not feasible, our study demonstrates stronger

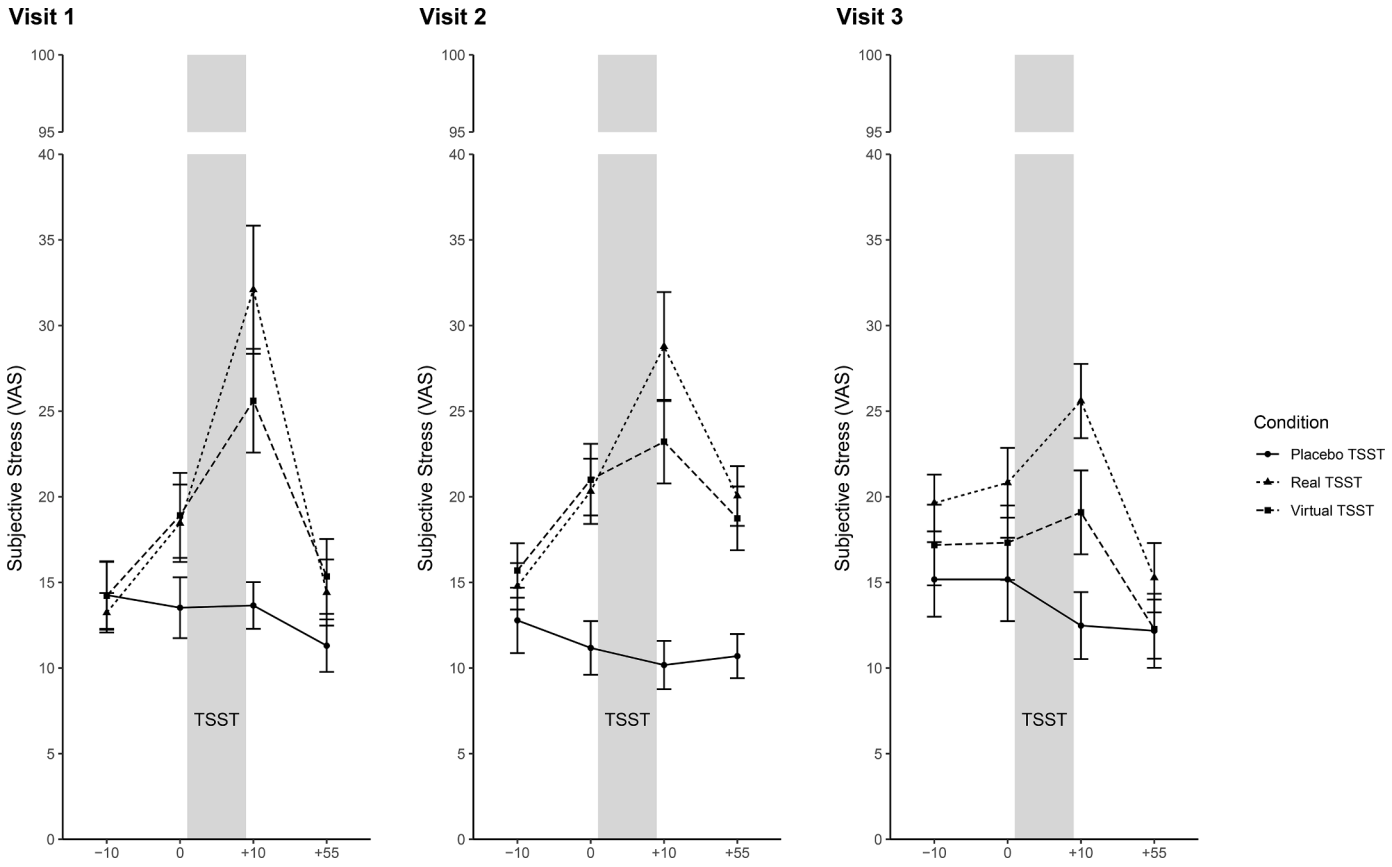

Fig. 5. Subjective Stress (C) of three conditions across the three visits (Mean ± SEM).

habituation of the HPA axis with a slightly lower level of initial HPA reactivity. In comparison, Kotlyar et al. [22], who also used an HMD, did not find any HPA reactivity. This, however, could be due to the fact that at that time, VR-glasses were quite heavy and impracticable [16].

Apart from technology, the virtual experience may also contribute to stress reactivity. While spatial presence, i.e., the sense of being there, has received some attention [29, 36], social presence in the context of VR-TSST has been rather underrepresented. Our findings suggest that – surprisingly – the degree of social presence does not seem to influence the stress response.

4.4. Strengths and limitations

One of the strengths of this study is the use of a placebo group, which allowed controlling for the effect of technology on stress reactivity. This is even more important since almost none of our participants were experienced with VR, and the technology could have contributed to their stress. Furthermore, we had a 2-person mixed-gender panel to single out gender effects. Yet, Goodman et al. [8] found that three jury members may possibly evoke stronger responses. Further, we do not investigate habituation effects over a longer period of time, according to Schommer et al. [30], who did this for a real-life TSST. Finally, none of the participants experienced adverse effects such as cybersickness, dizziness or headaches, in contrast to Montero-López et al. [26].

5. Conclusions

As suggested by prior research (Zimmer et al., 2019[35]), our results demonstrate that the VR-TSST is well applicable and has the potential to produce robust endocrine, autonomous, and self-reported stress responses. Yet, our findings also show that the VR-TSST is less applicable to repetitions, as there tends to be a stronger decrease in cortisol reactivity in repeated exposures. The replicability at shorter intervals seems problematic for the traditional TSST and especially for the VR-TSST. Participants show a habituation effect to the VR-TSST on the third visit regarding subjective stress and thus bears the risk that subjects may no longer see the task as a social challenge but as some form of entertainment. Furthermore, persons who are more experienced with VR or immersive technology may generally be prone to habituate more quickly in VR-based stressor tasks.

6. Author contributions

O.D.K. designed the study, O.D.K., J.X.K. and A.F. were involved in the recruitment of participants and acquisition of data, L.B. developed the virtual reality system, L.M.G. and R.P. performed the analysis of saliva cortisol and provided lab resources, O.D.K., B.P. and A.G. analyzed the data, O.D.K., A.G. and A.F. wrote the first draft of the manuscript, L.M.G., J.X.K., B.P., H.H., I.K.E. and R.P. contributed extensively to the first draft, A.G. prepared the figures and tables, I.K.E. and H.H. provided lab resources for assessments and supervised the study. All authors have approved the final manuscript.

Funding

Open access funding is provided by Medical University of Vienna.

CRediT authorship contribution statement

Oswald D. Kothgassner: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft. **Andreas Goreis:** Formal analysis, Methodology, Visualization, Writing – original draft. **Lisa M. Glenk:** Formal analysis, Writing – review & editing. **Johanna Xenia Kafka:** Data curation, Writing – review & editing. **Bettina**

Pfeffer: Formal analysis, Writing – review & editing. **Leon Beutl:** Software, Visualization. **Ilse Kryspin-Exner:** Resources, Supervision, Writing – review & editing. **Helmut Hlavacs:** Resources, Supervision, Writing – review & editing. **Rupert Palme:** Resources, Writing – review & editing. **Anna Felnhofer:** Data curation, Writing – original draft.

Declaration of Competing Interest

All authors have no conflict of interest to declare.

Acknowledgments

We want to thank Johanna Rudyk, Anna-Katharina Heinzle, Nora Poesl, and Farina Flick for their assistance in the study piloting and the acquisition of parts of the data.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.physbeh.2021.113618.

References

[1] J.N. Bailenson, J. Blascovich, A.C. Beall, J.M. Loomis, Interpersonal distance in immersive virtual environments, Pers. Soc. Psychol. Bull. 29 (7) (2003) 819–833, https://doi.org/10.1177/0146167203029007002.
[2] D. Bates, M. Mächler, B.M. Bolker, S.C. Walker, Fitting linear mixed-effects models using lme4, J. Stat. Softw. 67 (2015), https://doi.org/10.18637/jss.v067.i01.
[3] S.S. Dickerson, M.E. Kemeny, Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research, Psychol. Bull. 130 (3) (2004) 355–391, https://doi.org/10.1037/0033-2909.130.3.355.
[4] M.A. Fallon, M.M. Riem, L.E. Kunst, W.J. Kop, N. Kupper, Multi-modal responses to the Virtual Reality Trier Social Stress Test: a comparison with standard interpersonal and control conditions, Int. J. Psychophysiol. 161 (2021) 27–34, https://doi.org/10.1016/j.ijpsycho.2021.01.010.
[5] A. Felnhofer, H. Hlavacs, L. Beutl, I. Kryspin-Exner, O.D. Kothgassner, Physical presence, social presence, and anxiety in participants with social anxiety disorder during virtual cue exposure, Cyberpsychol. Behav. Soc. Netw. 22 (1) (2019) 46–50, https://doi.org/10.1089/cyber.2018.0221.
[6] Finch, W.H., Bolin, J.E., Kelley, K., 2019. Multilevel modeling using R. Multilevel Modeling Using Mplus.
[7] M.B. First, R.L. Spitzer, M. Gibbon, J.B.W. Williams, Structured Clinical Interview For DSM-IV Axis I Disorders, Clinician Version (SCID-CV), American Psychiatric Press, Washington, D.C., 1997.
[8] 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.
[9] M.R. Gunnar, B.M. Reid, B. Donzella, Z.R. Miller, S. Gardow, N.C. Tsakonas, K. E. Thomas, M. DeJoseph, J.J. Bendezú, Validation of an online version of the Trier Social Stress Test in a study of adolescents, Psychoneuroendocrinology 125 (2021), 105111, https://doi.org/10.1016/j.psyneuen.2020.105111.
[10] E.C. Helminen, M.L. Morton, Q. Wang, J.C. Felver, A meta-analysis of cortisol reactivity to the Trier Social Stress Test in virtual environments, Psychoneuroendocrinology 110 (2019), 104437, https://doi.org/10.1016/j.psyneuen.2019.104437.
[11] E.C. Helminen, M.L. Morton, Q. Wang, J.C. Felver, Stress reactivity to the Trier Social Stress Test in traditional and virtual environments: a meta-analytic comparison, Psychosom. Med. 83 (3) (2021) 200–211, https://doi.org/10.1097/PSY.0000000000000918.
[12] S. Het, N. Rohleder, D. Schoofs, C. Kirschbaum, O.T. Wolf, Neuroendocrine and psychometric evaluation of a placebo version of the 'Trier Social Stress Test, Psychoneuroendocrinology 34 (7) (2009) 1075–1086, https://doi.org/10.1016/j.psyneuen.2009.02.008.
[13] D.J. Hruschka, B.A. Kohrt, C.M. Worthman, Estimating between- and within-individual variation in cortisol levels using multilevel models, Psychoneuroendocrinology 30 (2005) 698–714, https://doi.org/10.1016/j.psyneuen.2005.03.002.
[14] P. Jönsson, M. Wallergård, K. Österberg, Å.M. Hansen, G. Johansson, B. Karlson, Cardiovascular and cortisol reactivity and habituation to a virtual reality version of the Trier Social Stress Test: a pilot study, Psychoneuroendocrinology 35 (9) (2010) 1397–1403, https://doi.org/10.1016/j.psyneuen.2010.04.003.
[15] R.P. Juster, B.S. McEwen, S.J. Lupien, Allostatic load biomarkers of chronic stress and impact on health and cognition, Neurosci. Biobehav. Rev. 35 (1) (2010) 2–16, https://doi.org/10.1016/j.neubiorev.2009.10.002.
[16] O. Kelly, K. Matheson, A. Martinez, Z. Merali, H. Anisman, Psychosocial stress evoked by a virtual audience: relation to neuroendocrine activity, Cyberpsychol. Behav. 10 (5) (2007) 655–662, https://doi.org/10.1089/cpb.2007.9973.
[17] C. Kirschbaum, B.M. Kudielka, J. Gaab, N.C. Schommer, D.H. Hellhammer, Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis, Psychosom. Med. 61 (2) (1999) 154–162.
[18] C. Kirschbaum, K.-.M. Pirke, D. 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.
[19] O.D. Kothgassner, A. Felnhofer, H. Hlavacs, L. Beutl, R. Palme, I. Kryspin-Exner, L. M. Glenk, Salivary cortisol and cardiovascular reactivity to a public speaking task in a virtual and real-life environment, Comput. Human Behav. 62 (2016) 124–135, https://doi.org/10.1016/j.chb.2016.03.081.
[20] O.D. Kothgassner, A. Goreis, L.M. Glenk, J.X. Kafka, L. Beutl, I. Krypin-Exner, H. Hlavacs, R. Palme, A Felnhofer, Virtual and real-life ostracism and its impact on a subsequent acute stressor, Physiol. Behav. 228 (2021), 113205.
[21] B.M. Kudielka, S. Wüst, Human models in acute and chronic stress: assessing determinants of individual hypothalamus–pituitary–adrenal axis activity and reactivity, Stress 13 (1) (2010) 1–14, https://doi.org/10.3109/10253890902874913.
[22] M. Kotlyar, C. Donahue, P. Thuras, M.G. Kushner, N. O'Gorman, E.A. Smith, D. E. Adson, Physiological response to a speech stressor presented in a virtual reality environment, Psychophysiology 45 (6) (2008) 1034–1037, https://doi.org/10.1111/j.1469-8986.2008.00690.x.
[23] M.R. Leary, A brief version of the fear of negative evaluation scale, Pers. Soc. Psychol. Bull. 9 (3) (1983) 371–375, https://doi.org/10.1177/0146167283093007.
[24] Magnusson, K., 2018. Powerlmm: power analysis for longitudinal multilevel models. https://github.com/rpsychologist/powerlmm/.
[25] B.S. McEwen, Stress, adaptation, and disease: allostasis and allostatic load, Ann. N. Y. Acad. Sci. 840 (1) (1998) 33–44, https://doi.org/10.1111/j.1749-6632.1998.tb09546.x.
[26] E. Montero-López, A. Santos-Ruiz, M.C. García-Ríos, R. Rodríguez-Blázquez, M. Pérez-García, M.I. Peralta-Ramírez, A virtual reality approach to the Trier Social Stress Test: contrasting two distinct protocols, Behav. Res. Methods 48 (1) (2016) 223–232, https://doi.org/10.3758/s13428-015-0565-4.
[27] R. Palme, E. Möstl, Measurement of cortisol metabolites in faeces of sheep as a parameter of cortisol concentration in blood, Z. Säugetierkd. Int. J. Mamm. Biol. 62 (2) (1997) 192–197. Suppl.
[28] R Core Team, R: A language and Environment For Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2019. Retrieved from, http://www.r-project.org/.
[29] A.S. Ruiz, M.I. Peralta-Ramirez, M.C. Garcia-Rios, M.A. Muñoz, N. Navarrete-Navarrete, A. Blazquez-Ortiz, Adaptation of the trier social stress test to virtual reality: psycho-physiological and neuroendocrine modulation, J. Cyber Ther. Rehabil. 3 (2010) 405–415.
[30] N.C. Schommer, D.H. Hellhammer, C. Kirschbaum, Dissociation between reactivity of the hypothalamus-pituitary-adrenal axis and the sympathetic-adrenal-medullary system to repeated psychosocial stress, Psychosom. Med. 65 (3) (2003) 450–460, https://doi.org/10.1097/01.PSY.0000035721.12441.17.
[31] Y. Shiban, J. Diemer, S. Brandl, R. Zack, A. Mühlberger, S. Wüst, Trier Social Stress Test in vivo and in virtual reality: dissociation of response domains, Int. J. Psychophysiol. 110 (2016) 47–55, https://doi.org/10.1016/j.ijpsycho.2016.10.008.
[32] M. Slater, S. Wilbur, A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments, Presence: Teleoperators Virtual Environ. 6 (6) (1997) 603–616, https://doi.org/10.1162/pres.1997.6.6.603.
[33] T.A.B. Snijders, R.J. Bosker, Multilevel analysis: An introduction to Basic and Advanced Multilevel Modeling (2nd ed.), Sage, London, 2011.
[34] C.C. Vatheuer, A. Vehlen, B. von Dawans, G. Domes, Gaze behavior is associated with the cortisol response to acute psychosocial stress in the virtual TSST, J. Neural Transm. (2021) 1–10, https://doi.org/10.1007/s00702-021-02344-w.
[35] 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.
[36] P. Zimmer, C.C. Wu, G. Domes, Same same but different? Replicating the real surroundings in a virtual Trier Social Stress Test (TSST-VR) does not enhance presence or the psychophysiological stress response, Physiol. Behav. 212 (2019), 112690, https://doi.org/10.1016/j.physbeh.2019.112690.