International Journal of Virtual Reality
Volume 19 Issue 2

IPI Press
ISSN: 1081-1451

Santl, J., Shiban, Y., Plab, A., Wüst, S., Kudielka, B. M., & Mühlberger, A. (2019). Gender Differences in Stress Responses during a Virtual Reality Trier Social Stress Test. International Journal of Virtual Reality, 19(2), 2-15. https://doi.org/10.20870/IJVR.2019.19.2.2912

Received: 04/2019

Published: 10/2019

Gender Differences in Stress Responses during a Virtual Reality Trier Social Stress Test

Julia Santl[1], Youssef Shiban[3], Andreas Plab[1], Stefan Wüst[2], Brigitte M. Kudielka[4], Andreas Mühlberger[1]

[1]Department of Psychology, Clinical Psychology and Psychotherapy, University of Regensburg, Regensburg, Germany
[2]Department of Psychology, Differential Psychology and Behavioral Genetic Research Group, University of Regensburg, Regensburg, Germany
[3]Department of Clinical Psychology, PFH- Private University of Applied Sciences, Göttingen, Germany
[4]Department of Psychology Medical Psychology, Psychological Diagnostics, and Research Methodology, University of Regensburg, Regensburg, Germany

Corresponding author: Andreas Mühlberger, andreas.muehlberger@psychologie.uni-regensburg.de
Key words: Virtual Reality – Trier Social Stress Test (TSST) – Skin Conductance Level – Heart Rate – Social Stress – Salivary Cortisol

Abstract

The Trier Social Stress Test (TSST) is commonly used to induce stress in laboratories by instructing participants to deliver a speech and to solve arithmetic tasks in front of a committee. Its implementation in virtual reality (VR) enables an investigation of stress responses under highly standardized controllable conditions.

The aim of this study was to compare stress responses among men and women in a VR version of the TSST (VR-TSST).

To this end, 16 women taking oral contraceptives and 16 men underwent the VR-TSST in a modified version including a competitor. Stress ratings, heart rate, electrodermal reactivity, and salivary cortisol responses were analyzed.

The VR-TSST induced endocrine, peripher-physiological and self-reported stress responses, indicated by a significant increase in heart rate, electrodermal activity and stress ratings as well as a small but significant cortisol response. Significant gender differences were found only for stress ratings. In conclusion, these findings confirm earlier results that VR is suitable to induce social stress both in males and females.

Introduction

The Trier Social Stress Test (TSST) developed by Kirschbaum et al. in 1993 is a highly standardized stress protocol that combines social evaluative threat as well as uncontrollability of the test situation (Kirschbaum et al., 1993). The test is introduced as an interview for a desired job. During the test, a participant is asked to give a self-presentation speech and thereafter to do a mental arithmetic task, both in front of a committee. This committee (usually two or three experimental collaborators) is instructed to keep a neutral facial expression. The participant believes that the complete performance is videotaped, however, usually either no real video recording is made or the videotape is not analyzed further. Typical results are the increase of heart rate, blood pressure, cortisol levels, perceived stress levels, and negative affect (Kirschbaum et al., 1993; Al'Absi et al., 1997; Boesch et al., 2014; Goodman et al., 2017; Hawn et al., 2015; Kelly et al., 2008; Kudielka et al., 2007; Kudielka et al., 2009). The TSST has been shown to cause two-fold increased cortisol levels and an increased heart rate by up to 25 beats per minute (bpm) in most participants (Kirschbaum et al., 1993; Kudielka et al., 2007; Dickerson and Kemeny, 2004). A major issue in the usage of the TSST is the required logistic effort (e.g., the need to organize an experimenter and a committee of at least two people per participant). Recently, virtual reality TSST paradigms have emerged with the aim of reducing the involved logistic effort as well as the the variability in participant/committee interactions whilst increasing the reliability of the TSST (Allen et al., 2017). The first study integrating the TSST into VR using a head mounted display (HMD) compared endocrine responses evoked by a TSST in front of a real versus a virtual audience and the authors found lower cortisol responses compared to the original version (Kelly et al., 2007). Furthermore, Kotlyar et al. (2008) could not find significant cortisol responses in a speech task in VR, but they did find significant cardiovascular responses (blood pressure and heart rate). Ruiz et al. (2010) reported an average increase in skin conductance levels (SCL) in all study participants and an increase in cortisol levels in most of them during a VR-TSST. Jönsson et al. (2010) assessed cortisol and cardiovascular reactivity parameters as well as their habituation with two TSST sessions in VR using a fully immersive Cave Automatic Virtual Environment (CAVE) system. Mean cortisol levels where significant in the first session but habituated in the second session while heart rate increases were comparable in both sessions. Both, mean cortisol habituation and dissociated stress habituation patterns between the SAM and the HPA response systems to repeated TSST provocations are well-known findings (Schommer et al., 2003; Wüst et al., 2005). Wallergaard et al. (2011) also used a CAVE-system to examine the VR-TSST (Wallergaard et al., 2011). Heart rates showed an average increase that was comparable to the conventional TSST (24-34%, (Kirschbaum et al., 1993). Fich et al. (2014) examined the design of the virtual environment as a further possible influencing factor of the stress reaction in the VR-TSST. During the VR-TSST and the following recovery phase participants in a virtual room without window showed a higher cortisol response compared to participants in an virtual room including windows. Such differences were not found regarding high frequency heart rate variability, heart rate, and t-wave amplitude. Shiban et al. (2016) compared the TSST in vivo to the TSST in VR, using a HMD, in order to examine whether and to what extent a virtual, female competitor influences the stress reaction. The female competitor completed the tasks perfectly just before the test participant performed the tasks. Skin conductance, heart rate, salivary cortisol, stress ratings, and cognitive appraisal were measured. For heart rate and stress ratings, a significant increase occurred during the TSST for all groups. For cortisol the in vivo group showed significantly higher increases in response to TSST than the two VR groups.

In various studies, self-report, physiological, and humoral responses to the TSST differed between females and males: For instance, ACTH and salivary cortisol levels in response to the TSST are found to be higher for men than for women, while in woman menstrual cycle and intake of oral contraceptives exert important effects on the HPA responsiveness, too (Kirschbaum et al., 1999; Kirschbaum et al., 1992a; Montero-Lopez et al., 2018). For men, even the sole expectation of psychosocial stress causes a significant cortisol response whereas women did not show this merely anticipatory reaction (Kirschbaum et al., 1992b). Furthermore, Kudielka et al. (1998) observed that elderly men compared to women showed significantly higher ACTH and cortisol responses while the elderly women more often stated to be stressed. No gender differences emerged regarding the parameters heart rate, adrenaline, and noradrenaline. By contrast Kelly et al. (2008) found neither cortisol levels nor heart rate to differ between men and women. Women, however, showed a higher subjective fear response as well as more irritability, confusion, and less positive affect right after, but not 45 minutes after the TSST.

Regarding VR-TSST, Montero-Lopez et al. (2016) compared the role of gender in response to two different versions. Men and women in the first condition received the VR-TSST on a screen while the other group completed the VR-TSST with a head-mounted display (HMD). Cortisol and SCL increased in both conditions, with SCL being higher in the HMD group. An interaction of gender and cortisol reactions was not observed. A noteworthy limitation of this study is that neither the women's cycle phase nor the intake of contraceptives was controlled despite their known influence on the HPA responsiveness (Kudielka et al., 2009; Kirschbaum et al., 1999).

The first aim of the present study was to confirm that the competitor enhanced VR-TSST (see (Shiban et al., 2016)) provokes substantial self-reported stress as well as endocrine and peripher-physiological stress responses in both woman and man. Secondly, we aimed to investigate whether there are gender differences by comparing endocrine, peripher-physiological, and self-reported stress among men and women. Due to previous studies on the original TSST, we expected women to report higher stress ratings and men to show higher endocrine responses to the VR-TSST.

Method

Participants

We recruited 32 participants (16 females, 16 males) through advertisements at the University of Regensburg. Due to the influence of fluctuating endogenous sexual hormones on stress responses due to the female menstrual cycle (Kirschbaum et al., 1999; Montero-Lopez et al., 2018), only women taking oral contraceptives were included. Exclusion criteria were age younger than 18 or older than 55, current psychotherapeutic or pharmacological treatment, cardiovascular or neurological illnesses, and intake of medication containing glucocorticoids, such as cortisol. The German version of the Mini International Neuropsychiatric Interview (M.I.N.I.) by Ackenheil et al. (1999) was used to exclude participants suffering from mental illnesses. Furthermore, regular smokers (more than five cigarettes per day) were excluded to avoid influence by nicotine. As caffeine affects the HPA axis, participants were asked not to consume any beverages containing caffeine at least three hours before the experiment. Participants' alcohol consumption per week and potential incidences of addiction in their families were assessed as well. Due to the circadian rhythm of the HPA axis, all stress exposures were performed between 1 and 7 p.m., as described by Wüst et al. (2000).

On average, participants were 23.16 years old ($SD = 3.28$, range = 20-30), with men ($M = 24.50$, $SD = 3.52$, range = 20-30) being older than women ($M = 21.81$, $SD = 2.46$, range = 20-28); ($t(30) = -2.50$, $p < .05$, Cohen's $d = 0.64$). To account for possible age effects on results, age was added as a covariate for analyses on gender effects. On average, participants showed values of 15.25 ($SD = 8.52$) in the Social Phobia Inventory (SPIN), with women ($M = 17.69$, $SD = 8.32$) showing higher values then men ($M = 12.81$, $SD = 8.26$). However, this difference was not statistically significant ($t(30) = 1.66$, $p = .11$). The Primary Appraisal Secondary Appraisal (PASA) questionnaire showed significantly higher values for women compared to men in the subscales challenge (women $M = 4.66$, $SD = 0.68$; men $M = 4.09$, $SD = 0.62$, $t(30) = -2.44$, $p < .05$, $d = 0.87$) and control expectancy (women $M = 4.97$, $SD = 0.54$; men $M = 4.13$, $SD = 1.45$, $t(30) = -2.18$, $p < .05$, $d = 0.77$). Secondary scales primary appraisal and secondary appraisal as well as the tertiary scale stress index were not different between genders.

Materials

The head-mounted display "Oculus Rift Development Kit 2" (Oculus VR, Irvine, California, USA) with a diagonal field of view of 100°, a fast head tracking system, rapid refreshing rates and a high resolution (1920*1080 pixels) was used to promote immersion in VR. The virtual environment was created using the Steam Source Game Engine (Valve Corporation, Bellevue, Washington, USA). With the use of Cybersession (version 5.6.94, VTplus GmbH, Würzburg, Germany) the virtual environment was controlled and timed.

The virtual environment consisted of a sparsely furnished entrance and a testing room equipped with a desk for the committee, chairs, and a camera (Figure 1). There were four virtual agents in the testing room: a man and a woman in the committee, both wearing white coats, a female experimenter and a young male competitor. The committee members showed a neutral facial expression without emotions and as a response they occasionally nodded. The virtual experimenter was the first agent to talk to the participant by explaining the tasks and introducing the competitor. The woman in the committee gave instructions, asked questions, disclosed false arithmetic, and ended tasks, whereas the man in the committee kept silent. The competitor was the first one to perform the tasks and fulfilled them perfectly. The behavior of all virtual agents was standardized in frequency across test sessions for all participants and over all tasks. In order to reinforce the competition, the competitor received positive feedback, whereas the participant only obtained neutral feedback. The virtual agents' reactions to the participant were activated by the experimenter via keyboard.

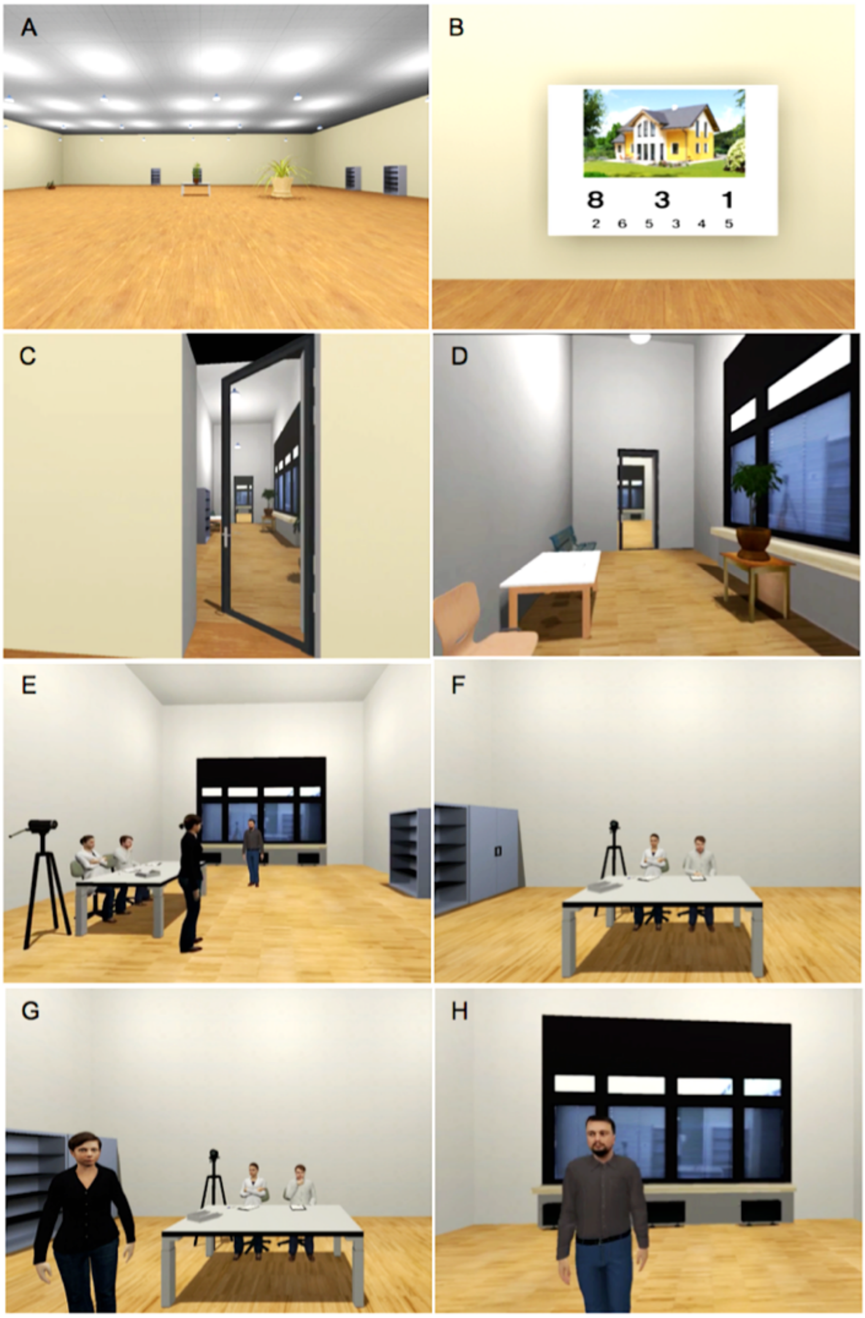

Figure 1: Virtual environment seen from the perspective of the participant.

A: Entrance with table in the middle, B: View into the hallway, C: testing room with all virtual agents, D: Committee, E: Virtual experimenter with committee in the background, F: Competitor.

Measures

Stress rating

The self-reported stress level was measured eight times: as a pre-stress level before the VR-TSST, at the end of the VR-TSST as well as 10, 20, 30, and 40 minutes after the end of the test. Furthermore, stress ratings for the speech and the arithmetic task were given retrospectively at the end of the VR-TSST. Stress ratings were obtained by asking participants to rate on a scale from 0 (no stress) to 100 (extreme stress) how stressed they felt at the moment.

Peripher-physiological measures
Electrocardiogram (ECG) and skin conductance level (SCL) were recorded continuously about 5 minutes before the beginning to ten minutes after the VR-TSST. For ECG, two surface electrodes (Ag/AgCl 40 mm) were attached to the sternum and to the left lower costal arch. For SCL, two surface electrodes (Ag/AgCl 8mm) with isotonic electrode gel were attached next to each other on the thenar eminence on the palm of the non-dominant hand.

Cortisol
Salivary cortisol was collected by means of cortisol salivettes (Sarstedt AG & Co., Nürnbrecht, Deutschland) at seven time points: at pre-stress (30 min before the VR-TSST), during the preparation (10 minutes before the VR-TSST), at the end of the TSST and 10, 20, 30, and 40 minutes after the end of the test. According to Dickerson and Kemeny (2004) peak levels of cortisol are reached between 21 and 40 minutes after the beginning of the stressful event. Therefore, we identified individual peaks for each participant within the time interval from time point end of TSST (Post 0) to time point +30 min. (Post 30).

iGroup Presence Questionnaire (iPQ)
The iPQ designed by Schubert (2003) investigates experience in computer environments and measures the feeling of presence in the virtual world. It consists of 14 items rated on a seven-point scale from -3 to +3 (fully disagree to fully agree). The iPQ consists of the three subscales spatial presence (5 items), involvement (4 items) and experienced realism (4 items) and one additional item ("sense of being there") not belonging to any of these subscales. Internal consistency of the iPQ is sufficient with Cronbach's α from .63 to .78 for the three components (Schubert, 2003).

Design
The independent variable of this quasi-experimental study consists of two gender groups: women and men. As repeated measurement factor, the factor "time" including 5 time points enables the investigation of temporal changes. Dependent variables include stress ratings, physiological reactivity assessed by EDA and electrocardiogram (ECG), as well as neuroendocrine responses assessed by salivary cortisol samples before, during, and after VR-TSST exposure.

Procedure
The exact procedure including all measures is depicted in the upper part of Figure 2.

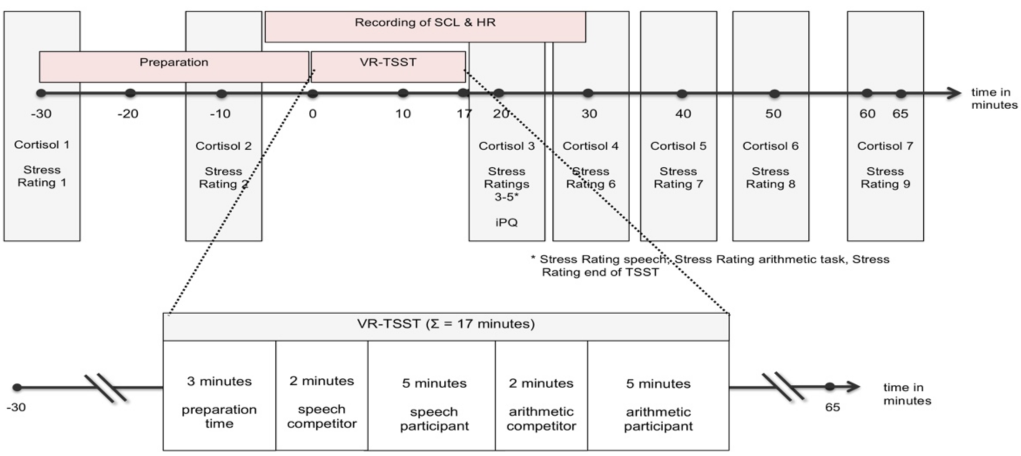

Figure 2: Flow chart of the experimental procedure.

Heart rate and SCL were recorded 5 minutes before to 10 minutes after the VR-TSST. The grey boxes show the chronology of measurements. Lower part: Detailed flow chart of the VR-TSST.

Having given written consent, participants were asked for a stress rating and the first salivary cortisol sample was collected to examine pre-stress cortisol levels. This was followed by a preparation phase of approximately 30 minutes, in the course of which the participant had to sit still and fill out questionnaires. During this preparation phase, another stress rating and another saliva cortisol sample was collected. After this preparation phase, there was a test phase of 17 minutes in VR, whereby participants found themselves in virtual reality five minutes before the start of the VR-TSST. The participants were asked to give a five-minute talk as part of an employment interview for their dream job. The participants believed that the performance is recorded. They had three minutes to prepare for the task. Directly after observing the competitor performing excellently and receiving positive feedback, it was their turn. The committee members behaved neutrally and told the participants to continue in case their speech ended before the five minutes were over. If there was a second pause, participants were inquired about potential negative traits and in case of another pause, they were asked about leadership qualification and personal suitability. After five minutes participants were interrupted and received neutral feedback. Afterwards, they watched the competitor solving the mental arithmetic task perfectly and receiving positive feedback once more. When it was their turn to do mental arithmetic for five minutes, they had to continuously subtract 17 from 2023 as fast and accurate as possible. In case of mistakes, they were told to start again, whereas in case of correct answers, they were instructed to complete the arithmetic task faster and to speak more clearly. If the participants gazed at the ceiling or the floor, they were instructed to look at the committee and in case of pronunciation mistakes to pronounce numbers correctly. The competitor, who was dealing with a different, equally difficult mental arithmetic task, often looked at the participant, making disapproving noises or laughing maliciously. Apart from the competitor and his behavior (Shiban et al., 2016) the procedure corresponded to the original version of the TSST (Kirschbaum et al., 1993). Feedbacks included oral statements as well as head movements and were initiated by the experimenter by keystrokes.
At the end of the test phase, another saliva sample was collected and another stress rating was given. Furthermore, stress ratings for the speech and the arithmetic task were given retrospectively. Moreover, participants filled in the iPQ. The following resting phase lasted approximately 45 minutes. Ten minutes after the end of the test, the electrodes were removed and recordings of physiological data ended. Saliva samples and stress ratings were collected three more times at 20, 30 and 45 minutes after the end of the test. The total test session lasted approximately two hours per participant.

Data reduction

Physiology
ECG and skin conductance level (SCL) were analyzed using the software Brain Vision Analyzer (version 1.20, Brain Products GmbH, Gilching, Germany). First, differences between the electrodes on the sternum and the costal arch were computed. A high-pass filter of 1.59 Hz (12db), a low-pass filter of 30 Hz (12db), and a bandpass filter of 50 Hz were applied to optimize R-wave detection. R-waves were identified using the R-wave detection solution in the Brain Vision Analyzer, controlled manually, and revised if necessary. Finally, heart rate was computed out of R-wave distances. For skin conductance level (SCL), the means for each of the seven segments were calculated. Raw mean values were first z- and then T-transformed for each participant over all segments to get values independent of the individual absolute level. Mean values of heart rate and electrodermal activity were further analyses for seven sequences: Pre-stress (3 min.), preparation (3 min.), competitor's speech (2 min.), subject's speech (5 min.), competitor's mental arithmetic (2 min.), subject's mental arithmetic (5 min.) and recovery (3 min.).

Cortisol
Saliva samples were stored at room temperature during the test and then put in the freezer at -20°C until their biochemical analysis. For biochemical analysis they were sent to the Department of Biological and Clinical Psychology at the University of Trier. After thawing for biochemical analysis, samples were centrifuged at 2000 g at 10° C for 10 minutes. Salivary cortisol concentrations were determined by a time resolved immunoassay with fluorescence detection (DELFIA) described elsewhere (Dressendörfer et al., 1992). The

assay sensitivity was 0.173 nmol/l. Intra- and inter-assay variabilities were less than 6.7% and 9.0%, respectively.

Statistical Analyses

Statistical analyses were performed using the statistics software Statistical Package for Social Sciences (SPSS version 24) (IBM, Armonk, New York, USA). The level of significance was tested two-sided for all analyses and accepted with $\alpha = .05$. Data were tested for normal distribution and variance homogeneity using the Kolmogorov-Smirnov- and Levene-Test prior to the analysis. However, normal distribution was presumed for further analysis due to the sample size greater than 30 (Bortz and Schuster, 2005). ANOVAs for repeated measurements with the within-subjects factor time (nine time points for the stress ratings, seven for SCL, HR, and cortisol) and the between-subjects factor gender (male, female) were calculated to analyze stress ratings, heart rate, SCL, and salivary cortisol. When variance homogeneity was violated, the Greenhouse-Geisser correction was applied. Partial eta square (η_p^2) was used to give effect size. 2 x 2 ANOVAs were calculated in case significant interactions resulted, in other cases where only main effects where significant post-hoc Bonferroni corrected joined *t*-tests were calculated. Although no significant main effect was found for gender in endocrine and peripher-physiological measures, in order to exploratively examine gender differences at specific time points *t*-tests for independent samples were calculated for each time of measurement. To detect differences between genders concerning age, *t*-tests for independent samples were calculated, additionally determining Cohen's *d*. The level of significance was tested two-sided for all analyses and accepted with $\alpha = .05$. Due to age differences between men and women, age was added as a covariate to the ANOVA.

Results

Stress rating

The mean stress level, which was determined nine times, before, during and after the VR-TSST, started from 20.41 ($SD = 22.69$) (on a scale from 0 to 100) at the beginning and increased within the VR-TSST to about 60.44 ($SD = 18.20$: see Figure 3). The repeated measurement ANOVA showed a significant main effect of time ($F(2.71, 81.54) = 81.54$, $p < .001$, $\eta_p^2 = .700$), and gender ($F(1,30) = 8.70$, $p < .01$, $\eta_p^2 = .225$), while gender x time interaction was not significant ($F(2.71, 81.54) < 1$, $p = .693$). Overall, women reported significantly higher stress than men. To follow up the main effect time, post-hoc Bonferroni corrected joined *t*-tests with a significance level of $\alpha = .006$ revealed the following significant differences between time points: The pre-stress rating was significantly lower than stress ratings given for speech and arithmetic ($ps < .001$), the stress rating during preparation was significantly lower than stress ratings given for speech ($p < .001$) and arithmetic ($p < .001$) and during recovery ($p = .002$). The mean stress level after the speech and arithmetic part did not differ from one another, both being significantly higher compared to all other measurement points (all $ps < .001$). The exploratory *t*-tests for differences between genders in specific time points, revealed lower values for men than for women at the time of speech ($t(30) = 2.15$, $p < .05$, $d = .07$), arithmetic ($t(30) = 2.27$, $p < .05$), end of TSST ($t(22.70) = 2.51$, $p < .05$), and 10 min. ($t(23.16) = 2.72$, $p < .05$), 20 min. ($t(21.29) = 2.10$, $p < .05$), and 30 min. after the end ($t(22.01) = 2.21$, $p < .05$).

A further analysis including age as a covariate was conducted in order to confirm that the effect of gender does not result from age differences between men and women. The interaction between time and age ($F(2.70, 78.30) < 1$, $p = .890$) and the main effect of the covariate age ($F(1,29) < 1$, $p = .376$) were not significant while the main effect of gender remained stable ($F(1, 29) = 9.29$, $p < .01$, $\eta_p^2 = .243$).

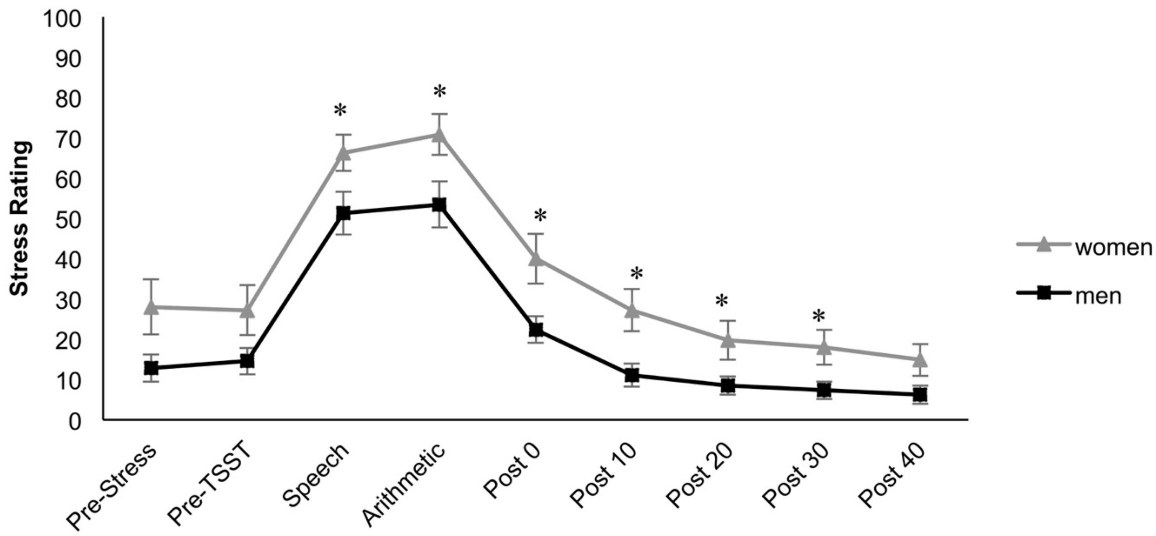

Figure 3: Subjective stress rating (± SEM) over the whole experiment in men and women.

*p < .05, asterisks refer to Bonferroni corrected post-hoc comparisons.

Heart rate

At the beginning, the average heart rate was 84.38 bpm (*SD* = 13.61) and increased to 100.17 bpm (*SD* = 17.92) during the stress test. This implies a mean increase of 18.87% resp. 15.79 bpm (Figure 4). The analysis of the heart rate showed a significant main effect of time ($F(3.04, 91.44) = 44.48$, $p < .001$, $\eta_p^2 = .597$). Neither the interaction gender and time ($F(3.04, 91.44) = 1.47$, $p = .227$) nor the main effect of gender ($F(1, 30) = 2.11$, $p = .157$) were significant. Post-hoc Bonferroni-corrected joined *t*-tests with a significance level of $\alpha = .007$ showed lower heart rate during pre-stress compared to all other measurement times except recovery (all *p's* < .001). Furthermore heart rate while preparation was lower than while speech competitor, speech, and recovery (all *p*'s < .01). Heart rate while own speech and competitor's speech was significantly higher than in all other measurement points ($p < .01$) except from own arithmetic. *t*-Tests for independent samples did not show significant gender differences at any measurement time.

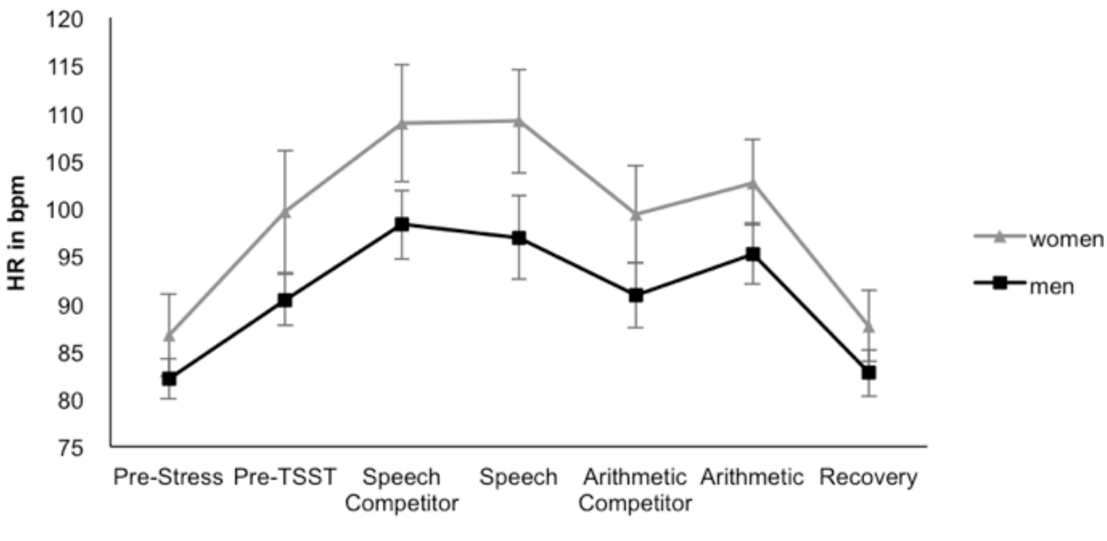

Figure 4: Mean heart rate responses (± SEM) to the VR-TSST in men and women.

3.3 Skin Conductance Level

SCL showed an increase of 22.77% (SD) during speech and 17.96% during the arithmetic task compared to pre-stress. The analysis of the SCL (Figure 5) showed a significant main effect of time ($F(2.91, 87.31) = 5.61$, $p < .01$, $\eta_p^2 = .158$). The interaction between gender and time ($F(2.91, 87.31) = 2.19$, $p = .097$) and the main effect of gender ($F(1, 30) < 1$, $p = 1$), were not significant. Post-hoc Bonferroni-corrected joined *t*-tests with a significance level of $\alpha = .007$ showed significant differences between the different time periods: SCL was significantly higher during preparation ($p = .003$) and own and competitor's speech than during pre-stress (*p*s = .001). No further significant time effect was found. *t*-tests for independent samples showed that men had a significantly lower SCL during competitor's arithmetic compared to women ($t(30) = -2.34$, $p < .05$).

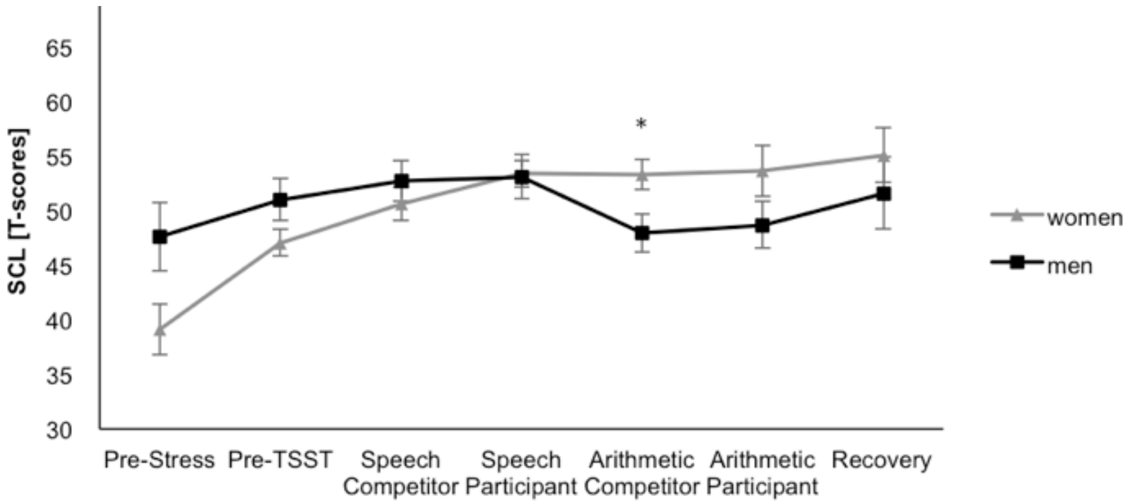

Figure 5 : Mean SCL responses (±SEM) in men and women during speech and arithmetic task of the competitor.

*p < .05, asterisks refer to Bonferroni corrected post-hoc comparisons.

Cortisol response

On average, participants had mean pre-stress cortisol levels of 4.66 nmol/l ($SD = 3.06$) and a mean individual peak level of 8.01 nmol/l ($SD = 5.94$), showing a mean increase of 3.35 nmol/l ($SD = 5.75$). Based on the study of Miller et al. (2013) an increase of 1.5 nmol/l from baseline to peak was defined as the criterion to separate between responders and non-responders. For responder rates, the difference between individual peak and baseline was calculated for every participant. Applying this criterion of Miller et al. to our study resulted in a responder rate of 50 percent (16 cortisol responders (8 female, 8 male), 16 non-responders (8 female, 8 male). The repeated-measures ANOVA with the repeated measurement factor time and gender as between subject factor (Figure 6) showed a significant main effect of time ($F(1.61, 46.70) = 4.35$, $p < .05$, $\eta_p^2 = .130$). Neither a significant interaction of gender and time ($F(1.61, 46.70) = 2.19$, $p = .133$) nor a significant main effect of gender ($F(1, 29) = 3.09$, $p = .089$) emerged. Post-hoc Bonferroni-corrected joined *t*-tests with a significance level of $\alpha = .007$ showed the following results: Cortisol levels at time point + 20 min. were significantly higher than those at timepoint + 40 min. ($p = .002$) and cortisol levels at timepoint + 10 min. were significantly higher than those at timepoint + 30 min ($p = .006$). Even if there was no significant main effect for gender, exploratory *t*-tests for every timepoint indicate that women had lower values than men at pre-stress ($t(30) = 2.76$, $p < .05$), preparation ($t(30) = 3.40$, $p < .01$) and right after the end of the stress test ($t(30) = 2.12$, $p < .05$). In sum, cortisol increased in response to the TSST, while no clear differences for gender could be confirmed.

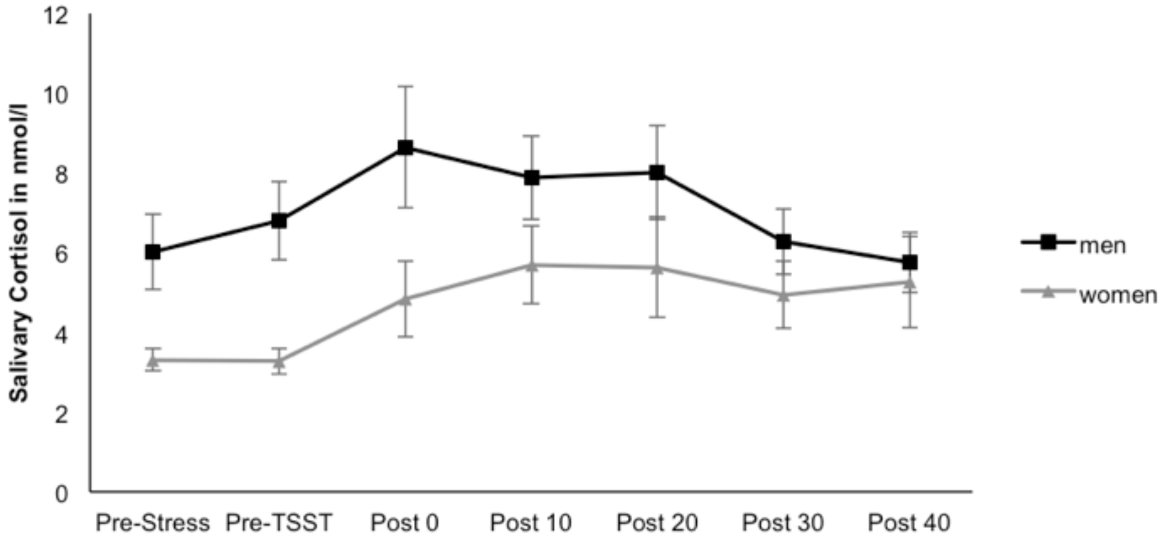

Figure 6: Mean cortisol responses (± SEM) to the VR-TSST in men and women.

iPQ

The presence questionnaire iPQ revealed high scores for the subscales “spatial presence” and “involvement” as well as average scores for subscale “experienced realism” and the additional item “sense of being there” (see Table 1). No significant differences between woman and men were found (all p’s > .44).

	woman	man	total
IPQ General P.	1.44 (1.26)	1.19 (1.60)	1.31 (1.42)
IPQ Spatial P.	2.30 (1.44)	2.00 (0.90)	2.15 (1.19)
IPQ Involvement	2.23 (1.45)	2.17 (1.26)	2.20 (1.34)
IPQ Realness	0.77 (1.17	0.45 (1.10)	0.61 (1.13)

Table 1 : Self-report of Presence

Note: Mean (SD) are given. IPQ = Igroup Presence Questionnaire, range -3 to 3.

Discussion

The objective of the present study was to confirm that the VR-TSST provokes substantial self-reported, endocrine, and peripher-physiological stress responses. Furthermore, we aimed to investigate whether there are gender differences by comparing stress responses between men and women.

Subjective stress level

The self-reported stress level increased constantly from the beginning and reached its peak at the arithmetic task, which showed a twofold increase. After the VR-TSST the stress dropped fast and the initial level was reached ten minutes after the stress test. The expected increase of the stress level throughout the TSST could be confirmed. In this study, the subjective stress level is the indicator with the highest increase from initial to peak level after the VR-TSST as well as the one with the highest observed effect sizes. Overall gender differences were observed as women claimed significantly higher stress levels than men which is in line with results of Kelly et al. (2008) and Kudielka et al. (1998). However, no differences between men and women in their stress response to the VR-TSST were found. Our results are in line with Hawn et al. who reported significant higher stress levels after VR-TSST compared to a neutral task in VR (Hawn et al., 2015). Furthermore, they did not find any differences in the stress level between the VR and traditional TSST. It must be considered that the measurements of the actual stress level during the speech and during arithmetic was given retrospectively after the VR-TSST to avoid disturbances for the participants due to the assessment during the test.

Peripher-physiological measurements

The heart rate showed an increase of 18.87% beats per minute from pre-stress to end of the stress test. This increase is slightly lower than the ones reported by Kirschbaum et al. (1993) and Kotlyar et al. (2008) for the original TSST (24 - 34% and 30% respectively). In their studies using a CAVE, Jönsson et al. (2010) reported an increase of 17% during verbal contribution and an increase of 22% during the arithmetic task, while Wallergaard et al. (2011) could only find an increase of 12% during the speech task and 9% during the arithmetic task. Using a HMD, Shiban et al. (2016) found an increase of 12% in the speech and 10% in the arithmetic task. Thus, our findings comparable to previous findings of the VR-TSST using HMD or CAVE. Heart rate did not vary significantly as a function of gender. This result is in accordance with the results of Kelly et al. (2008) for the standard-TSST.

SCL constantly increased during the TSST which is in line with the studies by Ruiz et al. and Montero-López et al. who reported an increase of the SCL from pre-stress and from preparation to the speech of the VR-TSST (Ruiz et al., 2010; Montero-Lopez et al., 2016). In our study there was an increase of 23% in verbal contribution and 18% in arithmetic from pre-stress level. For SCL there was a great effect of time in the repeated-measures ANOVA. Men and women did only differ significantly in their SCL responses during their competitor's arithmetic task, where women reached significantly higher values than men. In contrast to the heart rate, there was no significant recovery in SCL. A reason for this could be that the SCL should have been measured for a longer time period following the stress test. Furthermore, SCL is very susceptible to breathing and motion artifacts (Boucsein, 2012). During the recovery phase, relaxation after testing could have led to intensified breathing and sitting down as well as taking off glasses could have caused motion artifacts.

Cortisol

On average, there was a 72% increase in cortisol from pre-stress phase to the individual level of each participant. This is in line with Jönsson et al. (2010), that reported an increase of 88% for cortisol levels in the first session compared to baseline? Male participants showed higher cortisol levels at the first three measurement points compared to females which is in line with previous findings for the in vivo TSST (Kirschbaum et al., 1992a).

iPQ

The sense of presence experienced by the participants in our virtual environment was good with average to high ratings for the subscales involvement, spatial presence, and the general presence item ("sense of being there"), but only low ratings for the subscale experienced realism. We assume that realism can be enhanced in further experiments by using HDM with a higher field of view as well as resolution and to provide more movements of the virtual committee, Interestingly, however, the presence overall was sufficient to induce substantial response in the participants.

Summary

Taken together, our results confirm that the VR-TSST provokes substantial self-reported, endocrine, and peripher-physiological stress responses. Furthermore, our results show gender differences for some of the stress measurements: While self-reported stress levels were overall significantly higher in women, we could not confirm differences in stress responses. Physiological measures (heart rate and skin conductance level) did not discriminate between genders. The iPQ indicates that participants experienced a sense of presence in a virtual environment (no differences between genders). In general, we could see that the VR-TSST is an effective instrument for the induction of self-reported, physiological, and endocrine stress responses, being able to discriminate between genders in self-reported but not in physiological and endocrine measures.

Limitations

The sample was relatively small with overall 32 participants. Thus, especially the missing differences between genders may be due to low statistical power and should be replicated in larger samples. Furthermore, only young and healthy students were included. This reduces the generalization for status of age and the level of education. Besides this, men were somehow older than women. To account for this difference, age was used as a covariate in the case of significant gender differences.

Conclusion

To conclude, this study is further evidence that the TSST in VR is a suitable tool to produce stress in a laboratory setting for both men and women. This is shown in the rise of heart rate, skin conductance response, cortisol levels, and self-reported stress level. In line with previous studies on the in vivo TSST, women reported significantly higher levels of stress than men throughout the study, but showed similar self-reported, physiological and endocrine stress responses. The cortisol response in this VR-TSST study was lower than in traditional TSST studies, thus further work should enhance aspects of the virtual situation if cortisol response is an important dependent variable.

Acknowledgments

We would like to specially thank Andreas Ruider for the valuable support in designing and programming the virtual exposure scenario.

References

Ackenheil, Stotz-Ingenlath, Dietz-Bauer, Vossen, 1999. MINI Mini International Neuropsychiatric Interview, German Version 5.0. 0 DSM IV. *Munich: Psychiatric University Clinic*.

Al'Absi, Bongard, Buchanan, Pincomb, Licinio, Lovallo, 1997. Cardiovascular and Neuroendocrine Adjustment to Public Speaking and Mental Arithmetic Stressors. *Psychophysiology* 34 (3): 266–75.

Allen, Kennedy, Dockray, Cryan, Dinan, Clarke, 2017. The Trier Social Stress Test: Principles and Practice. *Neurobiology of Stress* 6: 113–26.

Boesch, Sefidan, Ehlert, Annen, Wyss, Steptoe, La Marca, 2014. Mood and Autonomic Responses to Repeated Exposure to the Trier Social Stress Test for Groups (TSST-G). *Psychoneuroendocrinology* 43: 41–51.

Bortz and Schuster, 2005. Statistik Für Sozial-Und Humanwissenschaftler. *Springer Medizin Verlag, Heidelberg* 4 (4.5): 4–6.

Boucsein, 2012. *Electrodermal Activity*. Springer Science & Business Media.

Dickerson and Kemeny, 2004. Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin* 130 (3): 355.

Dressendörfer, Kirschbaum, Rohde, Stahl, Strasburger, 1992. Synthesis of a Cortisol-Biotin Conjugate and Evaluation as a Tracer in an Immunoassay for Salivary Cortisol Measurement. *The Journal of Steroid Biochemistry and Molecular Biology* 43 (7): 683–92.

Fich, Jönsson, Kirkegaard, Wallergaard, Garde, Hansen, 2014. Can Architectural Design Alter the Physiological Reaction to Psychosocial Stress? A Virtual TSST Experiment. *Physiology & Behavior* 135: 91–97.

Goodman, Janson, Wolf, 2017. Meta-Analytical Assessment of the Effects of Protocol Variations on Cortisol Responses to the Trier Social Stress Test. *Psychoneuroendocrinology* 80: 26–35.

Hawn, Paul, Thomas, Miller, Amstadter, 2015. Stress Reactivity to an Electronic Version of the Trier Social Stress Test: A Pilot Study. *Frontiers in Psychology* 6: 724.

Jönsson, Wallergaard, Osterberg, Hansen, Johansson, Karlson, 2010. Cardiovascular and Cortisol Reactivity and Habituation to a Virtual Reality Version of the Trier Social Stress Test: A Pilot Study. *Psychoneuroendocrinology* 35 (9): 1397–1403.
Kelly, Tyrka, Anderson, Price, L Carpenter, 2008. Sex Differences in Emotional and Physiological Responses to the Trier Social Stress Test. *Journal of Behavior Therapy and Experimental Psychiatry* 39 (1): 87–98.
Kelly, Matheson, Martinez, Merali, Anisman, 2007. Psychosocial Stress Evoked by a Virtual Audience: Relation to Neuroendocrine Activity. *CyberPsychology & Behavior* 10 (5): 655–62.
Kirschbaum, Bartussek, Strasburger, 1992b. Cortisol Responses to Psychological Stress and Correlations with Personality Traits. *Personality and Individual Differences* 13 (12): 1353–57.
Kirschbaum, Kudielka, Gaab, Schommer, Hellhammer, 1999. Impact of Gender, Menstrual Cycle Phase, and Oral Contraceptives on the Activity of the Hypothalamus-Pituitary-Adrenal Axis. *Psychosomatic Medicine* 61 (2): 154–62.
Kirschbaum, Pirke, Hellhammer, 1993. The '91Trier Social Stress Test'92--a Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology* 28 (1–2): 76–81.
Kirschbaum, Wüst, Hellhammer, 1992a. Consistent Sex Differences in Cortisol Responses to Psychological Stress. *Psychosomatic Medicine* 54 (6): 648–57.
Kotlyar, Donahue, Thuras, Kushner, O'Gorman, Smith, Adson, 2008. Physiological Response to a Speech Stressor Presented in a Virtual Reality Environment. *Psychophysiology* 45 (6): 1034–37.
Kudielka, Hellhammer, Kirschbaum, Harmon-Jones, Winkielman, 2007. Ten Years of Research with the Trier Social Stress Test'97 revisited. *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior* 56: 83.
Kudielka, Hellhammer, Wüst, 2009. Why Do We Respond so Differently? Reviewing Determinants of Human Salivary Cortisol Responses to Challenge. *Psychoneuroendocrinology* 34 (1): 2–18.
Kudielka, Hellhammer, Hellhammer, Wolf, Pirke, Varadi, Pilz, Kirschbaum, 1998. Sex Differences in Endocrine and Psychological Responses to Psychosocial Stress in Healthy Elderly Subjects and the Impact of a 2-Week Dehydroepiandrosterone Treatment. *The Journal of Clinical Endocrinology & Metabolism* 83 (5): 1756–61.
Miller, Plessow, Kirschbaum, Stalder, 2013. Classification Criteria for Distinguishing Cortisol Responders from Nonresponders to Psychosocial Stress: Evaluation of Salivary Cortisol Pulse Detection in Panel Designs. *Psychosomatic Medicine* 75 (9): 832–40.
Montero-Lopez, Santos-Ruiz, Garcia-Rios, Rodriguez-Blazquez, Rogers, Peralta-Ramirez, 2018. The Relationship between the Menstrual Cycle and Cortisol Secretion: Daily and Stress-Invoked Cortisol Patterns. *International Journal of Psychophysiology* 131: 67–72.
Montero-Lopez, Santos-Ruiz, Garcia-Rios, Rodriguez-Blazquez, Perez-Garcia, Peralta-Ramirez, 2016. A Virtual Reality Approach to the Trier Social Stress Test: Contrasting Two Distinct Protocols. *Behavior Research Methods* 48 (1): 223–32.
Ruiz, Peralta-Ramirez, Garcia-Rios, Munoz, Navarrete-Navarrete, Blazquez-Ortiz, 2010. Adaptation of the Trier Social Stress Test to Virtual Reality: Psycho-Physiological and Neuroendocrine Modulation., *J. Cyber Ther. Rehabil* 3: 405–15.
Schommer, Hellhammer, Kirschbaum, 2003. Dissociation between Reactivity of the Hypothalamus-Pituitary-Adrenal Axis and the Sympathetic-Adrenal-Medullary System to Repeated Psychosocial Stress. *Psychosomatic Medicine* 65 (3): 450–60.
Schubert, 2003. The Sense of Presence in Virtual Environments: A Three-Component Scale Measuring Spatial Presence, Involvement, and Realness. *Zeitschrift Für Medienpsychologie* 15 (2): 69–71.
Shiban, Diemer, Brandl, Zack, Mühlberger, Wüst, 2016. Trier Social Stress Test in Vivo and in Virtual Reality: Dissociation of Response Domains. *International Journal of Psychophysiology* 110: 47–55.
Wallergaard, Jönsson, Johansson, Karlson, 2011. A Virtual Reality Version of the Trier Social Stress Test: A Pilot Study. *Presence* 20 (4): 325–36.

Wüst, Federenko, van Rossum, Koper, Hellhammer, 2005. Habituation of Cortisol Responses to Repeated Psychosocial Stress'97further Characterization and Impact of Genetic Factors. *Psychoneuroendocrinology* 30 (2): 199–211.
Wüst, Wolf, Hellhammer, Federenko, Schommer, Kirschbaum, 2000. The Cortisol Awakening Response-Normal Values and Confounds. *Noise and Health* 2 (7): 79.