Computers in Human Behavior 62 (2016) 124–135

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full length article

Salivary cortisol and cardiovascular reactivity to a public speaking task in a virtual and real-life environment

Oswald D. Kothgassner [a,*,1], Anna Felnhofer [b,1], Helmut Hlavacs [c], Leon Beutl [c], Rupert Palme [d], Ilse Kryspin-Exner [a], Lisa M. Glenk [e,f]

[a] Working Group Clinical Psychology and Health Psychology, University of Vienna, Austria
[b] Department of Pediatrics and Adolescent Medicine, Medical University of Vienna, Austria
[c] Research Group Entertainment Computing, University of Vienna, Austria
[d] Unit of Physiology, Pathophysiology and Experimental Endocrinology, University of Veterinary Medicine Vienna, Austria
[e] Karl Landsteiner Research Institute for Neurochemistry, Neuropharmacology, Neurorehabilitation and Pain Treatment, Landesklinikum Mauer, Austria
[f] Department of Comparative Medicine, Messerli Research Institute, University of Veterinary Medicine Vienna, Medical University Vienna and University Vienna, Austria

ARTICLE INFO

Article history:
Received 12 August 2015
Received in revised form
17 March 2016
Accepted 21 March 2016

Keywords:
Stress
Virtual reality
Salivary cortisol
Heart rate variability
Public speaking

ABSTRACT

Public speaking is a well-known psychosocial stress to occur in social-evaluative situations. This study examined self-reported, autonomic and endocrine stress responses to a 5-min public speaking task. Participants were asked to present either in front of i) a real audience, or ii) a virtual audience or iii) an empty virtual lecture hall. Thus, the main objective of this study was to examine the influence of real or virtual social stimuli on stress reactivity. Additionally, possible sex differences in stress responses were evaluated. Sixty-six women and men (20–33 years) underwent a multidimensional assessment of stress including self-reported state anxiety, heart rate, heart rate variability and saliva cortisol secretion. Results showed comparable increases in all stress responses in both the real and the virtual public speaking group. These findings indicate that the *Self Preservation Theory* is not limited to physically present social entities, but may also be extended to virtual social stimuli; as such this observation is also in line with the so called Media Equation Concept. Implications of the current results for therapy and research are subsequently discussed.

1. Introduction

Psychosocial stress is a widely recognized phenomenon which may occur in diverse social evaluative situations including for instance public speaking. When confronted with a socially threatening stimulus, two complementary mechanisms come into effect and shape the individual's acute response, the autonomic nervous system: the so called Sympatho-Adrenal-Medullary (SAM) axis and neuroendocrine system: the Hypothalamus-Pituitary-Adrenal (HPA) axis. Both axes "are instigated from the hypothalamus, where sensory inputs and serum-based feedback mechanisms monitor the level of environmental demand (i.e., stress) and the 'internal' state of the organism" (Bitsika, Sharpley, Sweeney, & McFarlane, 2014, p. 1). In other words, the two mechanisms help the organism adapt to environmental influences by initiating physiological as well as behavioral responses (e.g., the fight-or-flight response; c.f., Taylor et al., 2000).

While the SAM axis acts via the Sympathetic Nervous System (SNS) and is generally known to show a considerably fast reaction, the HPA axis is slower and mainly acts via the endocrine system. As the autonomic stress reaction via the SAM axis includes an increased blood flow and elevated heart rate as well as increased sweating and pupil dilation, SAM axis responses may best be assessed using measures such as heart rate or heart rate variability (HRV; Bitsika et al., 2014). HPA axis reactivity, in turn, is closely linked to the production and release of the stress hormone cortisol, which activates a feedback-loop to regulate stress-related physiological arousal. HPA axis activity may best be assessed using salivary cortisol as a biomarker as it reflects the amount of unbound, biologically active cortisol and may be sampled using non-invasive methods (Kirschbaum, Kudielka, Gaab, Schommer, &

* Corresponding author. Working Group Clinical Psychology and Health Psychology, University of Vienna, Liebiggasse 5/3, 1010, Vienna, Austria.
E-mail address: oswald.kothgassner@univie.ac.at (O.D. Kothgassner).
[1] Oswald D. Kothgassner and Anna Felnhofer contributed equally to this paper.

http://dx.doi.org/10.1016/j.chb.2016.03.081
0747-5632/

Hellhammer, 1999; Kudielka, Buske-Kirschbaum, Hellhammer, & Kirschbaum, 2004; Kudielka, Schommer, Hellhammer, & Kirschbaum, 2004; Schommer, Hellhammer, & Kirschbaum, 2003).

1.1. Social Self-Preservation Theory

A meta-analysis (Dickerson & Kemeny, 2004) has shown that if a situation or task includes social-evaluative elements as well as stimuli that are perceived as uncontrollable, cortisol changes are the largest and cortisol recovery rates the lowest. To explain this phenomenon, the authors make use of the so called *Social Self-Preservation Theory* (Dickerson & Kemeny, 2004). According to this theory, the HPA axis supports preservation of the social self by monitoring the environment for threats to the individual's self-esteem and social status; in case of a threat, it triggers self-evaluative cognitions and emotions and increases the release of cortisol in order to cope with the situation. Public speaking tasks, which are perceived as both uncontrollable and a social-evaluative threat because of the fear of failure and of being judged by others, have especially been considered as highly effective in evoking this chain of reactions (c.f. Dickerson & Kemeny, 2004).

The exposure to a social stressor has not only been shown to have a wide-ranging immediate impact on the individual's physiological response (c.f. Dickerson & Kemeny, 2004). It has also been discussed to be associated with cognitive and affective processes that may both directly and indirectly contribute to the development of various diseases and disorders such as depression (McEwen, 2005), cardiovascular diseases (Kemp, Quintana, Felmingham, Matthews, & Jelinek, 2012; McEwen, 1998) and immune dysfunction (Glaser & Kiecolt-Glaser, 2005). Furthermore, psychosocial stress has also been considered to be closely linked to the retention of anxiety disorders, obesity as well as obsessive compulsive disorders, OCD (Dishman et al., 2000; Pittig, Arch, Lam, & Craske, 2013).

1.2. Social stress and virtual reality

In light of its potent role in pathogenetic processes, the prevention or treatment of psychosocial stress has therefore increasingly been integrated in health promotion programs and in the treatment of psychological disorders. Traditionally, therapeutic exposition techniques are used to gradually decrease anxiety. Most commonly, they include confrontation with a feared stimulus in real-life (c.f., Wolpe, 1961). As an alternative to real-life exposure, virtual reality has increasingly come into use over the past decades (c.f., Krijn, Emmelkamp, Olafsson, & Biemond, 2004; Parsons & Rizzo, 2008; Powers & Emmelkamp, 2008). Virtual reality exposure therapy (VRET) provides a valuable intermediate step between imagination and exposure in real-life. Thus, VRET may be especially well-suited to patients who may have problems imagining a fearful stimulus or may not be ready to be exposed *in vivo*. Also, the therapist has maximal control over the stimulus as s/he may stop the simulation as soon as it becomes too overwhelming for the patient. Apart from the apparent gains of reducing therapy costs and time, VRET bears the additional advantage of increasing a patient's compliance and decreasing drop-out during exposure (Garcia-Palacios, Botella, Hoffman, & Fabregat, 2007; Repetto et al., 2011).

Among others, VRET has been developed to treat social anxiety: patients may learn new adaptive strategies in virtual social contexts and may then, in a next step, generalize these behaviors to a real-life social setting (Bordnick, Traylor, Carter, & Graap, 2012). Various studies support the assumption that virtual social scenarios may be used to evoke substantial levels of social anxiety on different levels, including cognitive, physiological or endocrine responses (e.g. Anderson et al., 2013; Garau, Slater, Pertaub, & Razzaque, 2005; Garcia-Lopez et al., 2006; Kampmann et al., 2016; Klinger et al., 2005; Krijn et al., 2004; Morina, Ijntema, Meyerbröker, & Emmelkamp, 2015; Opris et al., 2012; Owens & Beidel, 2015; Sarver, Beidel, & Spitalnick, 2014).

Also, virtual simulations may be used to help the patient train his/her social skills: Among existing programs are online social skills trainings (e.g. Lehenbauer, Kothgassner, Kryspin-Exner, & Stetina, 2013), virtual reality (VR), exposure applications (e.g., Gerardi, Cukor, Difede, Rizzo, & Rothbaum, 2010; Harris, Kemmerling, & North, 2002; Klinger et al., 2005) or even approaches which revert to existing virtual environments such as Massively Multiplayer Online Role-Playing Games (MMORPGs) or Second Life (e.g., Yuen et al., 2013). Their success – similarly to real-life exposure – relies on evoking social stress reactivity. Thus, it is crucial that those virtual cues which are supposed to represent social entities are perceived as such. According to evolutionary perspectives, the human brain is prepared to automatically respond to virtual entities in the same way it would to ‘real’ human beings – an assumption which has comprehensively been described in the *Media Equation* concept (Nass & Moon, 2000; Reeves & Nass, 1996).

Although some studies have already collected data in support of this theory (e.g., Kothgassner et al., 2014; Von der Pütten, Krämer, Gratch, & Kang, 2010), only few have picked up on this notion in the context of social-evaluative stress situations such as public speaking tasks. Previous investigations have examined the stress reaction to either a real audience (e.g., Kirschbaum et al., 1999; Kudielka, Hellhammer, & Wüst, 2009) or a virtual audience (e.g., Felnhofer et al., 2014; Jönsson et al., 2010; Slater, Pertaub, Barker, & Clark, 2006), but not to both. A comprehensive study comparing multiple measures (including self-report anxiety, physiological and endocrine responses) of social stress in a standardized experimental setting using a control group to control for specific influences of the used VR technology is still pending.

Among existing studies, some have shown that virtual audiences in a virtual public speaking task may provoke substantial levels of autonomic stress (as measured by heart rate) and self-reported anxiety (e.g. Felnhofer et al., 2014; Pertaub, Slater, & Barker, 2002; Slater et al. 2006). For instance, Hartanto et al. (2014) used three social scenarios, a neutral situation, a blind date and a job interview to evoke social stress and succeeded in achieving a substantial increase in heart rate (HR) and cognitive appraisal. Also, the authors showed that positive feedback received from virtual social entities leads to a reduction of anxiety and arousal and that negative feedback has a contrary effect. Similarly, Owens and Beidel (2015) found that a virtual environment may provoke significant increases of HR and skin conductance level (SCL) in comparison to a baseline condition and a control group. Yet another study (Villani, Repetto, Cipresso, & Riva, 2012) demonstrated heightened levels of SCL and self-reported anxiety in both a virtual and real-life job interview setting.

In most studies, however, stress indicators are restricted to measures of HR or SCL and subjective reports and do not account for cortisol as a crucial indicator for social-evaluative stress situations (c.f. Dickerson & Kemeny, 2004). Those studies which compare participants’ cortisol levels in similar social evaluative scenarios have produced inconsistent results: Kelly, Matheson, Martinez, Merali, and Anisman (2007) who introduced their participants to simulated job interviews found a significantly lower cortisol stress responses in the virtual group than in the real-life group. In contrast, Kotlyar et al. (2008) reported no change in cortisol concentrations during a speech task within a VR environment; and yet another study using a job interview showed both an increase in SCL and HPA (Montero-López et al., 2015). Finally, a study by Jönsson et al. (2010) which focused on the habituation of cortisol stress responses in a virtual stress test found an increase of salivary cortisol compared to the baseline only in the first exposure session and a habituation of stress responses in the second session. In sum, however, no study has to date compared stress responses on all levels and used a standardized experimental design when exposing participants to a real and virtual social stressor.

1.3. Sex, social stress and virtual reality

There is a definitive lack of knowledge about possible differences between men and women regarding both, their stress response to a social stimulus and their reaction to a virtual social entity. This is, however, all the more important, as differences in either one of the above mentioned domains would require sex-specific approaches to therapy. Past studies which exposed men and women to a real-life social-evaluative threat such as a public speaking task, found no sex differences in stress responses (Kirschbaum et al., 1999; Kudielka, Buske-Kirschbaum et al., 2004; Kudielka, Schommer et al., 2004; Kudielka, Hellhammer, & Wüst, 2009). Similarly, Kelly, Tyrka, Anderson, Price, and Carpenter (2008) did not find any variations in HPA reactivity between their male and female participants. However, the authors reported that women showed greater increases in negative affect, fear and confusion as well as decreases in happiness immediately following the social stress challenge. A game-based virtual simulation featuring competitive elements, social-evaluative threat and uncontrollability resulted in elevated cortisol levels and higher negative affect only in socially anxious males, not females (Brom et al., 2014). Similarly, Hemmeter et al. (2005) investigated a social evaluative threat in a virtual scenario and showed that males had higher cortisol responses than females. Montero-López et al. (2015) in turn, found no sex differences in stress reactivity, but instead reported differences in presence experiences between men and women: their result that women report to be more present in a stressful virtual situation, however, is in contrast to a study by Felnhofer, Kothgassner, Beutl, Hlavacs, and Kryspin-Exner (2012) who found men to have higher levels of presence in a virtual public speaking task.

Also, research has implied sex specific interactions with virtual entities. For instance, Szell and Thurner (2013) found that their female users valued interactions in virtual environments differently from their male counterparts; and in a study by Bailenson, Blascovich, Beall, and Loomis (2003), sex differences were detected with regard to mutual gaze behavior and keeping interpersonal distance to virtual avatars. Hence, in studies using virtual reality not only the stress response *per se* should be considered, but also the virtual entity itself in shaping an individual's stress reaction.

Furthermore, there is speculation that men and women might substantially differ in their physiological and behavioral response to a social stressor. Although all individuals show a (physiologically bound) fight-or-flight reaction when acutely threatened, Taylor et al. (2000) propose that females in contrast to males show a behavior described as "tend and befriend" – tending entails activities to promote safety and protect oneself as well as the offspring, whereas befriending is aimed at creating and maintaining social networks. Similarly, a study found women to show greater cortisol changes in response to social rejection stressors, whereas men were more responsive to achievement challenges (Stroud, Salovey, & Epel, 2002).

1.4. Objectives

Given that most virtual scenarios are validated against real-life scenarios using only one or a maximum of two stress indicators (mostly HR and self-reported measures), the current study set out to provide a more comprehensive and inclusive approach to validate a virtual social stimulus. By examining HR, heart rate variability (HRV) and self-reported anxiety measures and salivary cortisol, this study accounts for key indicators of a social-evaluative stress response (c.f., Dickerson & Kemeny, 2004). Thus, it contributes to the further advancement of a field which to date has been characterized by contradictory findings (e.g. Jönsson et al., 2010; Kelly et al., 2007; Kotlyar et al., 2008; Montero-López et al., 2015). The corresponding research question (RQ) is the following:

RQ 1: Are there differences regarding stress responses between participants speaking in front of a real audience and participants presenting in front of a virtual audience?

The verification of the *Social Self-Preservation Theory* (Dickerson & Kemeny, 2004) and the *Media Equation* concept (Reeves & Nass, 1996; Nass & Moon, 2000) requires a careful experimental setup to single out the influence of social cues among other non-social stressors. As VR technology *per se* may provoke a significant amount of stress it is, thus, crucial to control for its influence (c.f., Hartanto et al., 2014). By introducing an additional group to control for the virtual scenario, the participant's motor reactions or the apparatus used to immerse participants, this study allows for an evaluation of the impact of virtual social stimuli on the stress response (c.f., Kothgassner et al., 2014; Von der Pütten et al., 2010).

RQ 2: Are there differences regarding stress responses between the two groups speaking in front of an audience compared to a control group speaking in an empty virtual lecture hall?

Research focusing on sex differences in virtual scenarios using social stressors is challenged by the fact that these observed differences may be attributed to three sex specific responses: (1) stress reactivity in social situations (e.g., social rejection vs. social evaluation, Stroud et al., 2002), (2) experiences of virtual environments (Felnhofer et al., 2012; Montero-López et al., 2015), and (3) reactions to virtual social entities (Szell & Thurner, 2013). Hence, the current field of research on sex differences and VR is in need of carefully designed experiments which may disentangle these different influencing factors (Montero-López et al., 2015). This study provides a first approach to this subject, the according research questions are:

RQ 3: Are there differences between women and men regarding the stress response to the public speaking task?

RQ 4: Is there a difference in stress response between women and men with respect to the experimental condition (VR vs. in-vivo vs. control)?

2. Methods

On the basis of the above mentioned research questions, a 3 (experimental groups) x 2 (sex) between subject design was chosen to compare participants' stress responses to (1) a real audience in a real lecture hall, (2) a virtual audience in a virtual lecture hall and (3) an empty virtual lecture hall. (1) The first experimental group was asked to present in front of a physically present audience of a standardized number of 20 people. The presentation took place in a real university lecture hall and participants stood on a little podium behind a speaker's desk with an integrated display to watch the standardized presentation slides. (2) The second experimental group was instructed to present in front of 20 virtual characters in a virtual lecture hall which was a simulation of the real university lecture hall. (3) The control group presented in the same virtual lecture hall, but the lecture hall in this condition did not contain any audience in order to control for the influence of social stimuli. The groups in the virtual lecture hall (2 and 3) were connected to a virtual reality system while they stood in a 3 × 3 meter laboratory room behind a speaker's desk which they were also able to see in the virtual environment. Similar to the real lecture hall, participants in the virtual environment were able to track their presentation slides on a virtual display which was integrated into the virtual speaker's desk.

2.1. Procedure

The experiment was conducted at the University of Vienna VR-lab between March 2012 and June 2014; the experimental sessions started between 13.30 h and 15.30 h and took approximately 20 min excluding 30 min of pre-testing and 30 min of post-testing. The assessments were conducted according to a standardized protocol. Upon arrival to the laboratory, participants were escorted to a quiet, temperature-controlled room (23°-24 °C) where they were introduced to the course of the experiment and were handed the informed consent form. After signing it, participants were asked to fill out a demographic survey, a short interview to assess inclusion criteria (physical and mental illness) as well as psychological questionnaires. Furthermore, relevant anthropomorphic measures (e.g. height, weight) were recorded. Following this, the experimental procedure consisting of four periods with 5 min duration each started:

2.1.1. Waiting period

Prior to instruction, all participants were guided to the laboratory where they were asked to wait at a heightened table and remain in an upright position during the whole experiment.

2.1.2. Preparation period

All participants were asked to give a 5 min speech about Bhutan and its geography, culture and governmental form. Prior to the speech they had 5 min to get acquainted with the 20 slides of their presentation.

2.1.3. Public Speaking task (Stressor)

Participants were randomly assigned to one of the three experimental groups. Depending on the group, participants were guided to either the real lecture hall or the laboratory. After a short instruction (and the attachment of the VR equipment for the Groups 2 and 3), the participants started their presentations.

2.1.4. Recovery period

After the public speaking task the participants had 5 min to recover in the waiting room and to fill out a psychological questionnaire.

The current study was conducted according to the Declaration of Helsinki, and thus a comprehensive informed consent was provided to all participants prior to their voluntary participation. All participants were allowed to withdraw from the experiment at any time.

2.2. Participants

A total of 66 healthy women (n = 33) and men (n = 33) aged from 20 to 33 years (mean age = 24.17, SD = 3.05) participated in the current study. Female and male participants were equally balanced over the three experimental groups (n = 22 per group). All participants were students who received a course credit for their participation. None of the female participants administered any oral contraceptives and all were tested in the luteal phase of their menstrual cycle in order to ensure a better comparability of the females' endocrine stress responses to those of the male participants (Kirschbaum et al., 1999). An inclusion criterion for participation was an age ranging between 20 and 35 years to ensure comparability of physiological indicators (see Bigger et al., 1995; Kudielka, Buske-Kirschbaum et al., 2004; Kudielka, Schommer et al., 2004) and exclusion criteria were medication, smoking and a diagnosed medical or mental illness. Additionally, participants were instructed to not ingest caffeine or alcohol in the morning and not to eat for at least six hours prior to the start of the experiment. Table 1 shows means (M) and standard deviations (SD) for demographics. There were no significant differences between the three experimental conditions (groups) or sexes regarding their experience as a speaker assessed as the number of presentations held over the period of the last 5 years (GROUP: $F(2,60) = 0.824$; $p = 0.443$; $par\ \eta^2 = 0.027$; SEX: $F(1,60) = 1.515$.; $p = 0.108$; $par\ \eta^2 = 0.042$) and the experience with computer systems measured on a 5-point Likert-scale (“no skills” – “expert skills”) (GROUP: $F(2,60) = 0.536$; $p = 0.588$; $par\ \eta^2 = 0.018$; SEX: $F(1,60) = 1.796$; $p = 0.189$; $par\ \eta^2 = 0.029$). Furthermore, groups did not differ regarding the participants’ body mass index (BMI) ($F(2,60) = 1.348$; $p = 0.268$; $par\ \eta^2 = 0.043$). Yet, resulting from the lower BMI values of the female participants there was a sex difference $F(1,60) = 8.337$; $p = 0.005$; $par\ \eta^2 = 0.122$) over all groups.

2.3. Apparatus and virtual reality environment

Participants were immersed into Virtual Reality (VR) using a head mounted display (eMagin Z800 3D, Bellevue, Washington) with an integrated head-tracking unit (360° horizontal, >60° vertical) and two stereoscopic SVGA OLED displays (resolution: 800 × 600 triad pixels). A F710 Wireless Gamepad (Logitech Europe S.A., Switzerland) was used by the experimental groups 2 and 3 to flip the virtual presentation slides, while group 1 used a M185 Wireless Mouse (Logitech Europe S.A., Switzerland).

The virtual simulation was developed using Visual Studio C++ Express as well as the Ogre3D graphics engine (http://www.ogre3d.org/) in combination with Quick GUI. The virtual lecture hall was modeled according to a typical university lecture hall (see Fig. 1). All 3D-models and real-time rendering were achieved using Blender 3D; the textures were processed with GIMP (http://www.gimp.org/). To create human faces and mimics, photographs from the front and the side were loaded into Blender3D. The virtual audience showed a standardized minimum of emotions during the speech, to keep it comparable to the behavior of the real audience. Audio files from the real lecture hall were used in the virtual scenario to create

Table 1
Means and standard deviations of control factors over all three groups.

Group	Women		Men	
	M	SD	M	SD
Experience as speaker				
Real audience	2.7	0.65	3.2	1.17
Virtual audience	2.6	0.67	2.7	0.65
Control	2.6	0.82	2.9	0.30
Computer experience				
Real audience	3.3	1.42	3.6	1.04
Virtual audience	3.5	1.21	3.9	0.94
Control	3.2	1.17	3.6	0.69
Body Mass Index (BMI)				
Real audience	20.9	2.40	21.3	1.97
Virtual audience	20.8	1.36	22.7	1.40
Control	21.1	1.93	22.9	1.93

Fig. 1. Virtual Reality environment (demo-version) was used with a standardized number of 20 virtual listeners: i) Birdseye view on the lecture hall ii) virtual lecture hall first person point-of-view iii) point-of-view of the participants.

an adequate sound environment for the virtual lecture hall; in the empty virtual lecture hall only sizzling sounds could be heard at irregular intervals. Participants were able to turn their heads but they could not move freely around the virtual lecture hall.

2.4. Psychological measures

The *Personal Report of Confidence as a Speaker* (PRCS; Paul, 1966) is a widely used assessment for fear of public speaking. The 12-item questionnaire version of Hook et al. (2008) was used in the current study to analyze possible differences in the groups at baseline level. The 12 PRCS items (e.g. "I am fearful and tense all the while I am speaking before a group of people") are rated on a 4-point-Likert-scale ("does apply" – "does not apply"). Internal consistency of the measurement was high in this study (Cronbach's $\alpha = 0.87$).

The *State Trait Inventory* (State Subscale; STAI-S; Spielberger, Gorsuch, & Lushene, 1970) was used to assess the situational anxiety level of each participant prior to and shortly after the public speaking task. The STAI-S consists of 20 items (e.g. "I'm nervous") which were rated on 4-point-Likert-scale ("very" – "not at all"). Internal consistency of the measurement in the current study was high in the pre-assessment (Cronbach's $\alpha = 0.87$) as well as in the post-assessment (Cronbach's $\alpha = 0.88$).

2.5. ECG measures

The electrocardiogram (ECG) was monitored via three electrodes connected to a portable M-EXG (*Schuhfried*, *Mödling*, Austria) linked via Bluetooth (10 mW) to a personal computer. ECG was recorded at 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 side of the body to measure heart activity. A common ground electrode was placed on the neck of the participants. Electrodes were attached at the beginning of the waiting period to record heart rate (HR) and heart rate variability (HRV) through all periods of the experimental session (20 min). HR reflects the stress-related sympathetic activity (SAM axis) and was monitored at 5 ms intervals through the speech. Data were computed to 60s intervals using KUBIOS HRV software kit (Biosignal Analysis and Medical Imaging Group, Finland). High HR values indicate more beats per minute (bpm) and high physiological arousal, whereas low HR values reflect less bpm and lower physiological stress. The root mean square of successive differences (rMSSD) was obtained in accordance with the recommendations of the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996). The rMMSD is a time-domain measure reflecting a short-time measure of HRV. In this study, it was used to detect predominantly changes in the parasympathetic tone (SAM axis) during the experiment. Log-transformed rMSSD values were calculated from the beat-to-beat intervals for four 5 min-periods according to the experimental periods. High log-transformed rMSSD values indicate low physiological arousal while low values are interpreted as higher physiological arousal.

2.6. Salivary cortisol measures

To absorb saliva, commercial cotton swabs (Salivette®, *Sarstedt*, *Wiener Neudorf*, Austria) without any saliva-stimulating additives were used. Participants were thoroughly instructed how to collect their saliva by putting a swab into the cheek pouch, letting it saturate with saliva for approximately 60–80s and replacing the swab in the device container. The collected material was immediately stored at −20 °C. Prior to analysis, samples were thawed on ice and centrifuged at room temperature at 3000 g for 15 min to obtain the clear saliva. Sample processing and analyses were carried out with a highly sensitive cortisol enzyme immunoassay as described in detail by Palme and Möstl (1997) and previously applied in humans (Brom et al., 2014). A total of 10 µl of a 1:10 clear saliva dilution were used and all samples were assayed in duplicates. If duplicates showed coefficients of variance >10%, the assay was replicated and considered in the analysis only when a coefficient of variance <10% was achieved. If the sample volume fell below the limit needed to run duplicates or ran out before reaching a coefficient of variance <10%, the sample was dismissed from the analysis. Average intra- and inter-assay coefficients of variance were less than 10% and 15%, respectively. Salivary cortisol reflects the activity of the HPA axis; while high levels of cortisol in the human saliva indicate a high stress response, low levels are interpreted as a low response to a stressor. The first saliva samples were taken about 30 min after arrival in the laboratory and during the waiting period, the second saliva sample was taken 20 min after the beginning of the speech period and the third and last saliva sample was taken 20 min after the end of the recovery period.

3. Results

Statistical analyses were conducted using SPSS Version 20 (IBM, Corp. Armonk, USA), considering a significance level of $p < 0.05$ (two-tailed). Univariate ANOVAs and repeated measures ANOVAs with GROUP and SEX as between subject factors and TIME as a repeated factor as well as Bonferroni post hoc analyses were conducted to answer the four research questions. Data were tested for homogeneity of variance using the Levene Test. Greenhous-Geisser adjustments or the Huynh-Feldt correction were addressed to consider violations of the sphericity assumption (Girden, 1992). Partial Eta^2 (*par* η^2) was used as an effect size estimate for all ANOVA results. Areas under the response curves were computed using trapezoid formula (Pruessner, Kirschbaum, Meinlschmid, & Hellhammer, 2003) with respect to the increase of the stress response (AUCi). Salivary Cortisol AUCis were calculated over all three sample points (40 min); log-transformed rMSSD AUCis were computed for the experimental period (20 min). Table 2 displays all increases of stress response before (baseline measures) and during the public speaking task in percent (%) for direct comparisons and additional descriptive information to the AUCi calculation.

3.1. Psychological measures

3.1.1. Fear of public speaking

To control the effect of general fear of public speaking, an ANOVA was conducted to investigate baseline differences between the three groups (real audience, virtual audience, control group/no audience). There were no significant differences regarding the PRCS self-report questionnaire measuring general fear of public speaking between the three groups (GROUP: $F(2,60) = 0.256$; $p = 0.774$; $par\ \eta^2 = 0.009$). Moreover, no significant differences were found between female and male participants regarding fear of public speaking (SEX: $F(1,60) = 0.259$; $p = 0.589$; $par\ \eta^2 = 0.005$).

3.1.2. State anxiety

ANOVA calculations did not indicate any baseline differences between GROUP ($F(2,60 = 0.824$; $p = 0.443$; $par\ \eta^2 = 0.027$) and SEX ($F(1,60 = 0.230$; $p = 0.634$; $par\ \eta^2 = 0.004$) concerning the self-reported anxiety level (STAI-S) prior to the stressor. For the further analyses STAI-S ratings were computed with respect to increase (STAI-S pre-post differences). Univariate ANOVA revealed significant effects of GROUP ($F(2,60 = 5.766$; $p = 0.005$; $par\ \eta^2 = 0.161$), but again there was no significant influence of SEX ($F(1,60 = 1.110$; $p = 0.296$; $par\ \eta^2 = 0.018$) or any interaction effect GROUP x SEX ($F(2,60 = 0.207$; $p = 0.813$; $par\ \eta^2 = 0.007$). As indicated by the results above, the two experimental groups presenting either in front of a real (p = 0.009) or a virtual audience (p = 0.024), showed significantly increases in self-reported anxiety levels in comparison to the control group after the public speaking task. However, there was no significant difference in self-reported anxiety between the two experimental groups (p = n.s.).

Table 3 shows all means and standard deviations of the psychological measures.

3.2. ECG measures

Neither HR assessed for the first minute of the waiting period ($F(2,60 = 0.038$; $p = 0.963$; $par\ \eta^2 = 0.001$) nor the log rMSSD computed over the 5 min of the waiting period ($F(2,60 = 0.100$; $p = 0.905$; $par\ \eta^2 = 0.003$) showed significant group differences at baseline level. However, log rMSSD data showed no significant effect ($F(1,60 = 2.428$; $p = 0.124$; $par\ \eta^2 = 0.039$) for SEX, but there was a significant baseline difference regarding SEX for the HR recordings ($F(1,60 = 5.078$; $p = 0.028$; $par\ \eta^2 = 0.078$).

Table 2
Increases/Decreases of stress response in per-cent (women/men).

Group	%	(Women/men)
State anxiety		
Real audience	+15%	(+18%/+12%)
Virtual audience	+13%	(+18%/+9%)
Control	+1%	(+1%/+1%)
HR[a]		
Real audience	+22%	(+24%/+19%)
Virtual audience	+20%	(+24%/+15%)
Control	+8%	(+9%/+8%)
log rMSSD[a]		
Real audience	−11%	(−12%/−10%)
Virtual audience	−13%	(−16%/−10%)
Control	−3%	(−2%/−4%)
Salivary cortisol[b]		
Real audience	+47%	(+35%/+57%)
Virtual audience	+43%	(+36%/+51%)
Control	−12%	(+1%/−21%)

Notes:
[a] HR and log rMSSD increases were computed using baseline measures from the waiting period and the stress measures during the speech.
[b] Salivary Cortisol increases were computed using baseline measures −10 min before and stress measure +20 min after the speech.

Table 3
Means and standard deviations of the psychological measures for anxiety.

Group	Women				Men			
	Before stressor		After stressor		Before stressor		After stressor	
	M	SD	M	SD	M	SD	M	SD
PRCS 12								
Real audience	19.3	5.35	–	–	23.5	6.56	–	–
Virtual audience	20.3	4.36	–	–	21.4	7.80	–	–
Control	21.5	6.80	–	–	18.6	5.20	–	–
STAI-S								
Real audience	37.3	11.08	43.9	12.28	40.2	9.49	45.1	10.94
Virtual audience	34.8	7.33	41.2	7.97	36.5	5.09	39.6	7.39
Control	38.9	8.41	39.5	6.59	37.4	8.25	37.6	7.34

3.2.1. Heart rate

HR recordings were Greenhouse-Geisser adjusted according to results of the Mauchly Test for sphericity ($\chi^2 = 0.39.374$; $p < 0.001$; $\epsilon = 0.712$). To statistically compare the four experimental periods, HR recordings were computed in 5 min intervals. HR data for each minute over all four periods are displayed in Fig. 2. There was a significant increase over the four experimental periods (TIME: $F(2.136,128.153) = 154.030$; $p < 0.001$; $par\ \eta^2 = 0.720$) as well as a significant GROUP x TIME effect ($F(4.272,128.153) = 6.987$; $p < 04001$; $par\ \eta^2 = 0.189$), but no SEX x TIME effect ($F(2.136,128.153) = 1.906$; $p = 0.150$; $par\ \eta^2 = 0.031$) or any TIME x GROUP × SEX interaction effect ($F(4.272,128.153) = 0.932$; $p < 0.452$; $par\ \eta^2 = 0.030$). Participants did not show significant differences in HR assessed in 5 min intervals during the first waiting period (−10 min) or the preparation period (−5 min) (all ps = n.s.); however, all participants showed a significant HR increase during the public speaking task and during the recovery period (ps < 0.001). There was a higher increase in HR in the two experimental groups than in the control group (ps < 0.001). Bonferroni post hoc analyses reveal that the group presenting in front of a real audience showed significantly higher sympathetic arousal (p = 0.012) than the group in the control condition. When compared to the control group, the group presenting in front of a virtual audience only showed differences with a tendency to significance in case of an increasing threshold of $\alpha = 0.10$. There were no significant differences between the two experimental groups (p = n.s.).

3.2.2. Heart rate variability

Mauchly Test of sphericity indicated that the assumption of sphericity was violated for the log rMSSD calculation by repeated measures ANOVA ($\chi^2 = 25.756$; $p < 0.001$; $\epsilon = 0.921$). To correct this, the Huynh-Feldt correction was applied. On the one hand log rMSSD representing a HRV measure showed significant main effects regarding the decrease of parasympathetic tone (TIME: $F(2.762,165.72) = 38.893$; $p < 0.001$; $par\ \eta^2 = 0.393$) as well as statistically relevant differences over the three groups (GROUP x TIME: $F(5.524,165.72) = 3.700$; $p = 0.002$; $par\ \eta^2 = 0.110$). Nevertheless, post-hoc analyses indicate no significant group differences for the public speaking period (ps = n.s.). On the other hand neither a main effect of SEX x TIME ($F(2.762,165.72) = 38.893$; p=0.624; $par\ \eta^2 = 0.009$) nor an interaction effect between TIME x GROUP x SEX ($F(2.762,165.72) = 0.912$; $p = 0.482$; $par\ \eta^2 = 0.029$) can be

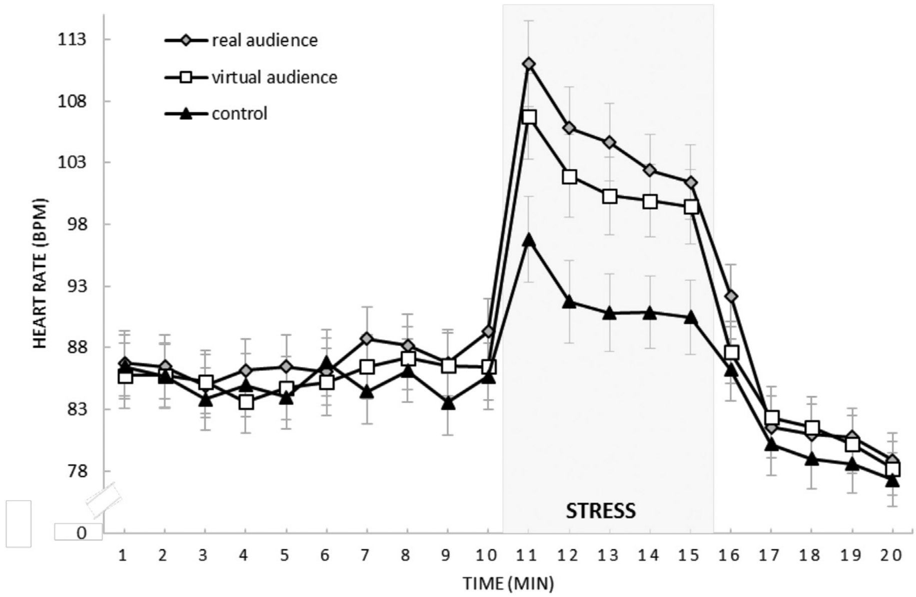

Fig. 2. Heart rate recordings in bpm (mean ± SEM) over the course of different experimental phases.

reported. According to the results of the AUCi calculations for the log rMSSD, the data reveal a significant main effect of GROUP (F(2,60 = 4.369; p = 0.017; $par\ \eta^2$ = 0.127) regarding the parasympathetic activity. As depicted in Fig. 3, both experimental groups showed lower parasympathetic activity than the control group, which should be interpreted as a higher stress response. However, Bonferroni post hoc analyses indicated no differences between the group presenting in front of the real audience and the control group (p = n.s.), but displayed significant differences between the control group and the group presenting in front of the virtual audience (p = 0.019). Yet, SEX had no statistically relevant influence on the participant's parasympathetic tone (F(1,60 = 0.602; p = 0.441; $par\ \eta^2$ = 0.010). Also, no interaction effect GROUP x SEX was found for the HRV measure (F(2,60 = 1.219; p = 0.303; $par\ \eta^2$ = 0.039).

3.3. Salivary cortisol measures

ANOVA revealed that neither GROUP nor SEX showed significant differences in salivary cortisol baseline levels prior to introducing the stressor (GROUP: F(2,60) = 0.182; p = 0.834; $par\ \eta^2$ = 0.006; SEX: F(1,60) = 0.292; p = 0.591; $par\ \eta^2$ = 0.005). Overall analysis for the salivary cortisol measurement was applied using a repeated measures ANOVA; it revealed a significant main effect of the stress response over the three salivary cortisol measurement time points (TIME; F(2,120) = 11.668; $p < 0.001$; $par\ \eta^2$ = 0.163). In comparison to the baseline saliva cortisol measures (−10 min), a notable cortisol increase to the public speech task for the experimental groups was found in the saliva sample after the stressor onset (+20 min, p < 0.001) as well as for the recovery period (+30 min, p < 0.004). Regarding this increase of cortisol levels in saliva, a significant main effect of the GROUP x TIME (F(4,120) = 4.247; p = 0.003; $par\ \eta^2$ = 0.124) was shown, whereas no significant main effect of SEX x TIME (F(2,120) = 1.690; p = 0.189; $par\ \eta^2$ = 0.027) could be found in the current study. Measures of cortisol showed no significant differences between the experimental groups during exposure to the stressor (p = n.s.), yet the control group presenting in an empty virtual lecture hall showed a significantly lower salivary cortisol stress increase than the groups presenting either in front of a real audience (p = 0.021) or a virtual audience (p = 0.032). Moreover, there was no TIME x GROUP × SEX interaction effect (F(4120 = 1.622; p = 0.173; $par\ \eta^2$ = 0.051). Additionally, results of the AUCi calculations from the salivary cortisol data suggest a significant main effect of GROUP (F(2,60) = 7.462; p = 0.001; $par\ \eta^2$ = 0.199), showing differences between the control group and the real audience (p = 0.003) as well as virtual-audience-group (p = 0.006), but no statistically relevant differences between both experimental groups (p = n.s.) according to the results of the repeated measures ANOVA. Furthermore, the analyses reveal no significant effect of SEX (F(1,60 = 0.075; p = 0.785; $par\ \eta^2$ = 0.001) as well as no significant interaction effect GROUP x SEX (F(2,60 = 1.962; p = 0.148; $par\ \eta^2$ = 0.062). Fig. 4 shows all salivary cortisol data (means ± SEM) for all groups and both sexes.

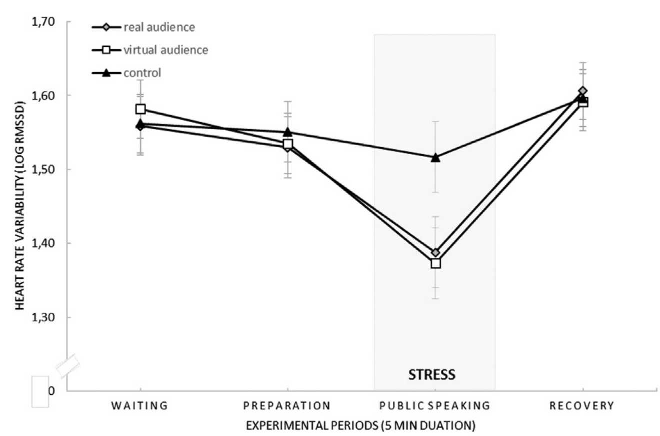

Fig. 3. Heart rate variability assessed by the log rMSSD (mean ± SEM) for all experimental periods (each 5 min). Data were computed to 5 min intervals.

4. Discussion

Given that most virtual scenarios are validated against real-life scenarios using only one or a maximum of two stress indicators, the current study set out to provide a more holistic approach. Thus, a multi-dimensional assessment of stress reactions, including stress reactivity of the HPA- and SAM-axis as well as self-report measures of anxiety, was used. By considering salivary cortisol as a crucial indicator of social-evaluative stress, we aimed to provide additional evidence for the validity of virtual social stressors. A public speaking task was chosen as it is regarded to be highly successful in eliciting psychosocial stress due to its uncontrollable nature and its inherent social-evaluative threat (c.f. Dickens & Kemeny, 2004). Also, this study was designed to verify the claims of both the *Social Self-Preservation Theory* (Dickerson & Kemeny, 2004) and the *Media Equation* concept (Nass & Moon, 2000; Reeves & Nass, 1996). It singled out the impact of social cues and controlled for influences of non-social stimuli (i.e. the VR technology) on the stress response by introducing a group presenting in front of an empty auditorium.

Finally, this study made an effort to investigate sex specific stress responses. A review of the corresponding literature suggests differences between men and women in their stress response to social situations (Stroud et al., 2002; Taylor et al., 2000), in their experiences of virtual environments (Felnhofer et al., 2012; Montero-López et al., 2015) and in their reaction to virtual social characters (Bailenson et al., 2003; Szell & Thurner, 2013). Thus, an experimental design was chosen which would allow us to disentangle these sex specific reactions. The corresponding results are discussed in more detail below.

4.1. Social stress reactivity

In this study, the two groups presenting in front of a virtual or real audience had higher overall levels of stress than the control group. Both experimental groups reported a rise in state anxiety levels (STAI-S) during their speeches and both showed significant increases in cardiovascular activity (SAM axis) as well as salivary cortisol concentrations (HPA activity). Also, there was no difference in all stress measures between the groups presenting in front of a virtual and a real audience but a large difference between these two groups and the controls. These findings support the assumption that the observed stress responses are a result of the confrontation with a social stimulus and not with the virtual apparatus. For instance, as HR reflects an ancestral behavioral activation system that occurs when the organism has to prepare to fight or flight, the significant increase of HR after the stressor onset may be attributed to the socially threating nature of the public speaking task. Also, when comparing the experimental VR group to the control group presenting in front of an empty virtual audience, the group presenting in front of a virtual audience showed significantly higher stress levels. This direct comparison of the two VR conditions further supports the assumption that it is the social stimulus which increases stress responses during the speech. In sum, these results help explain why prior studies using social virtual environments were successful in evoking specific social stress responses (c.f., Morina et al., 2015; Opris et al., 2012). Below, our findings shall be discussed in more detail and compared to current research.

In our study, self-reported anxiety levels did not differ between the two experimental groups, a finding that is perfectly in line with prior research (e.g. Anderson et al., 2013; Klinger et al. 2005; Krijn et al., 2004). This research also found participants to report comparable anxiety levels irrespective of whether they were confronted with a virtual or a real anxiety inducing stimulus. Similarly, Kampmann et al. (2016) found self-reported social anxiety to be the same in an *in vivo* and a VRET group and to differ significantly from a control group. However, the authors also reported general post-exposure anxiety levels to be significantly higher in the *in vivo* group than in the waiting-list controls or the VR group (with no post-exposure difference between controls and VR). Another study by Owens and Beidel (2015) seems to confirm this finding. In their intervention study the authors succeeded in demonstrating that affective appraisal of stress was significantly higher for the *in vivo* group than for the VR group.

In an attempt to explain the disparities between these studies and our experiment, one may first argue that while we examined healthy adults, both Kampmann et al. (2016) and Owens and Beidel (2015) conducted treatment studies with social anxiety disorder (SAD) patients. Prior research has shown that individuals with SAD – in contrast to healthy individuals – tend to overestimate their anxiety reaction and to revert to maladaptive coping strategies such as rumination (Hofmann, 2007). This may explain why these studies (e.g. Kampmann et al., 2016) resulted in significant post-exposure difference between the virtual and the real-life groups and why our study failed to find such differences. Future research focusing on disorder-specific coping strategies and subjective perceptions of stress could shed more light on this issue and further inform the usefulness of VR for therapy purposes.

In the present experiment, the SAM axis produced a stress response in the two presenting groups but not in the control group. However, the increase in the cardiovascular reactivity was not very distinctive. Only when the statistical threshold was set to a higher level, both experimental groups showed a significantly higher HR than the controls. This is generally in line with a study by Hartanto et al. (2014) who demonstrated that a job interview may trigger a stronger stress response than other social stress scenarios. However, even though their paradigm represented a social evaluative stressor like our public speaking scenario, the authors specifically focused on a particular type of social stress, one that is evoked by social interaction and especially by negative feedback from others. In the current study, however, there was no direct feedback from the audience. The virtual and real-life listeners followed a strict protocol which prohibited reactions other than neutral facial expressions. Thus, the comparability between the two studies may be diminished. Generally, our findings are supported by current models of social anxiety (Chen, Ehlers, Clark, & Mansell, 2002; Mansell, Clark, Ehlers, & Chen, 1999) which assume that an individual with SAD is subject to an attentional bias away from positive and negative facial expressions under social-evaluative threats. This may explain why this study was so successful in provoking social stress. Other research on physiological indicators of stress in social-evaluative situations shows results that are similar to our findings. Villani et al. (2012), for instance, reported a rise in SCL in a virtual and real-life job interview, Owens and Beidel (2015) showed

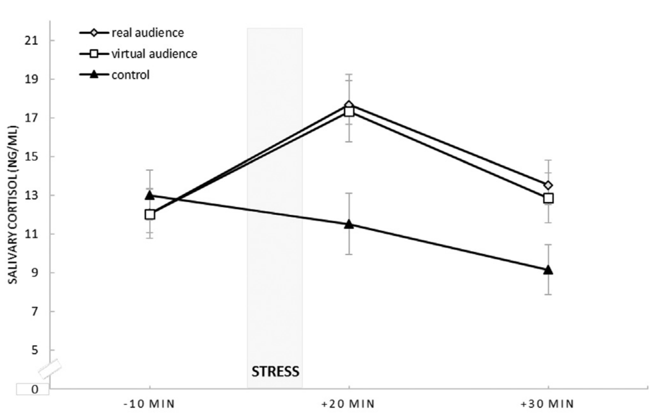

Fig. 4. Salivary Cortisol (mean ± SEM) for i) a baseline measure 10 min before the stressor, ii) a measurement of the stressor 20 min after the beginning of the stressor and iii) a measure of recovery cortisol habituation 30 min after the beginning of the stressor and 25 min after the beginning of the recovery.

an increase in SCL and HR in both real-life and in virtual exposure groups, and Slater et al. (2006) demonstrated a significant difference in HR between a virtual group and controls. However, looking at the setups one notices differences in the characteristics of the scenarios used. While our participants were asked to present in front of a full audience, Owens and Beidel (2015) as well as Slater et al. (2006) exposed their participants only to five audience members, and the participants in the study by Villani et al. (2012) were only exposed to one social entity. Future research should consider the question whether the number of audience members or interview partners may influence the intensity of the stress reaction. Also, studies should focus on the issue of whether the portrayed emotional expressions and gestures (c.f. Kang, Brinkman, van Riemsdijk, & Neerincx, 2016) or the type of interaction (direct negative or positive feedback, social-evaluative threats etc.) have a stronger impact on the social stress response.

With regards to HPA reactivity, both groups presenting either in front of a real or a virtual audience showed significant increases in salivary cortisol concentrations when compared to the control group. At the same time, the two experimental groups did not differ in their response. Generally, results from the current study correspond with prior research on VR exposure to social stressors (Jönsson et al., 2010; Montero-López et al., 2015). Jönsson et al. (2010) found an 88% gain of cortisol concentrations in their male participants during their virtual social stress task; in comparison, our study showed a 47% (real) and 43% (virtual) cortisol increase with no difference between the two groups. The almost double increase in the study by Jönsson et al. (2010) may be due to the fact that the authors implemented a highly immersive CAVE environment, while we used a conventional head mounted display. Research (e.g., Montero-López et al., 2015) comparing different technologies such as desktop-screens vs. goggles points towards an impact of mode of presentation on physiological reactivity during stress tasks. The more immersive a technology, the more likely it is that a person experiences presence in a virtual environment (Cummings & Bailenson, 2015). Presence, in turn, is a precondition for an emotional response; and a hallmark of presence is a reaction (both emotional and behavioral) that is congruent with one expected in a comparable real-life situation (c.f. Felnhofer et al., 2014). Hence, the highly immersive technology used by Jönsson et al. (2010) could be responsible for the considerable cortisol increase. Also, it is possible that the stimulus itself, a highly stressful job interview task with a standardized experimental protocol (Trier Social Stress Test; TSST, Kudielka, Buske-Kirschbaum et al., 2004; Kudielka, Schommer et al., 2004), may have produced stronger stress reactions than our public speaking task.

In sum, the hypothesis that real and virtual social stimuli may evoke a similar endocrine stress reaction is clearly supported in the current study. Because the control group lacked a comparable stress response during the speech, it is safe to assume that the significant increase in salivary cortisol in both experimental groups may be a result of the confrontation with a social stimulus. This is in line with the work of Dickerson and Kemeny (2004) who found that a situation which is perceived as a social threat by the individual may evoke the highest cortisol increases. Their *Social Self-Preservation Theory* (Dickerson & Kemeny, 2004) states that in situations in which an individual is in danger of being negatively evaluated by others and thus, faces embarrassment and detriment to his/her social self and self-esteem, the HPA axis comes into effect by triggering self-evaluative cognitions and emotions. All these mechanisms eventually help the individual to handle the threatening situation. Interestingly though, this self-preservative stress response seems not only to account for physically present social entities. As could be demonstrated in this study, virtual characters may also elicit a substantial cortisol response. This observation is in line with previous findings showing that individuals tend to engage in social behaviors with computers and to attribute a social entity to a machine when interacting with it (c.f. Bente, Krämer, & Petersen, 2002).

The *Media Equation Concept* (Reeves & Nass, 1996; Nass & Moon, 2000) explains why we are prepared to react to non-human stimuli – such as virtual characters – as if they were self-conscious social entities. According to this approach, the human brain unconsciously and automatically responds to human beings and virtual characters in the same way. The precondition for a correspondence in reactions to real and virtual social stimuli is that the latter contain sufficient social cues. Hence, a social reaction to a virtual entity does not necessarily require it to wear a human face but only to display a number of social cues that point to its existence as a social entity (Nass & Moon, 2000). In the current virtual environment, interaction with the virtual characters was set to a minimum, as the participants could not move around freely in the lecture hall and were not engaged in a conversation with the avatars. Yet, the virtual characters were not static, they displayed the usual behaviors of an audience: they were leaning on tables, their heads in their hands, and emitting sounds such as quiet chatting and coughing, thus facilitating the attribution of social characteristics to them.

To date, the validity of the Media Equation Concept (Reeves & Nass, 1996; Nass & Moon, 2000) has been demonstrated predominantly in studies allowing direct interaction with virtual entities, i.e. a study conducting a virtual ball tossing game (Kothgassner et al., 2014) and a study using a virtual agent for conversation (Von der Pütten et al., 2010). However, the concept has not yet been validated with a social evaluation stressor. Hence, this study is the first to provide strong support for the assumption that the mere presence of virtual others who display characteristics of social entities may lead to considerable social responses such as social stress reactivity. Also, having controlled for influences of VR technology on stress reactivity, this study has an advantage over past studies. It is now clear that VR *per se* does not elicit a stress response and that the observed social stress is due to the social influence of virtual others. Given this evidence, this type of stimulus may have strong promise for both the treatment of social anxiety and the assessment of anxiety outcomes during VR treatment.

4.2. Sex differences

In the present study, no sex differences in self-reported and physiological stress reactivity could be detected. Neither did male and female participants exhibit different physiological arousal levels or different self-reported levels of anxiety nor did their stress reactions differ in relation to stimulus type represented by the experimental conditions (virtual vs. real-life vs. control). The absence of sex specific stress reactivity and recovery patterns largely corresponds with results from several studies (e.g., Kelly et al., 2008; Kirschbaum et al., 1999; Kudielka, Buske-Kirschbaum et al., 2004; Kudielka, Schommer et al., 2004; Kudielka, Hellhammer, & Wüst, 2009). Montero-López et al. (2015), for instance, also did not find any sex differences in HPA or SAM axis activity in a VR adaptation of the TSST. Hemmeter et al. (2005) however, reported higher cortisol reactivity for men than for women in their dynamic VR. The contradiction between Hemmeter's and colleague's findings and our results may be explained by the fact that our female participants were all in the luteal phase and reported no use of oral contraceptives. Past findings demonstrate the crucial influence of the cycle phase on free cortisol responses (Kudielka & Kirschbaum, 2005). Women in the luteal phase exhibit cortisol increases that are similar to those of men, whereas women in the follicular phase and women taking oral contraceptives show lower cortisol responses than their male counterparts.

In sum, our results demonstrate that neither men nor women are influenced by the technological apparatus or type of condition (virtual vs. real-life stressor). However, we may not conclude from our data that the experience of the virtual environment was the same for our male and female participants. Montero-López et al. (2015) also found no sex differences in stress reactivity, but reported significant differences in presence experiences between men and women. Recent research may explain the different results for stress reactivity and virtual experiences: Two studies (Felnhofer et al., 2014; 2015) directly correlated physiological indicators such as HR and SCL to presence and found the physiological response to be unrelated to self-reported presence experiences. The authors conclude that presence and stress responses are logically distinct and that presence rather constitutes a precondition for an emotion to be at all elicited by VR. Furthermore, both sexes reacted similar to the social stressor used in this study. Hence, it may be stated, that male and female participants did not differ in their perception of the virtual social characters; this finding differs from the ones reported by Bailenson et al. (2003) as well as Szell and Thurner (2013) who both demonstrated sex specific experiences with avatars. An explanation for these disparities may be the fact that in our study – contrary to the ones mentioned above – interaction with the virtual characters was set to a minimum; participants could not move around freely in the lecture hall and were not engaged in a conversation with the avatars. Thus, it is conceivable, that in future scenarios with more virtual interaction, sex differences may be more prominent; all the more if the type of interaction or social stressor is considered as well (Kudielka & Kirschbaum, 2005).

Women have been shown to have greater cortisol changes in response to social rejection stressors, whereas men are regarded as more responsive to social evaluative challenges (Stroud et al., 2002). However, we only investigated one specific social-evaluative scenario in our study. Thus, it is conceivable, that men and women might differ in their stress reactivity depending on the type of social interaction used, e.g. social rejection vs. social evaluative challenges (c.f. Stroud et al., 2002). Also, some authors (Taylor et al. 2000) propose that in situations involving interpersonal stress females especially tend to show a behavior described as "tend and befriend". Following this, a direct comparison of different scenarios will most likely shed more light on this issue. Future studies should carefully select and control the type of social stressor used in order to shed more light on possible sex-specific reactions. By transferring, for instance, the Cyberball-paradigm, a well researched social rejection stimulus (c.f. Williams & Jarvis, 2006), to virtual reality, sex specific interactions with virtual entities may be studied under more standardized and clear-cut circumstances. Some studies (e.g. Kassner, Wesselmann, Law, & Williams, 2012; Kothgassner et al., 2014) have already implemented this paradigm in VR, but have not explicitly explored differences between male and female reactions to being ostracized by a virtual character. A simulation containing different and clearly separable types of stressors (i.e. social-evaluative stressors like the TSST and interpersonal stressors like the Cyberball-Paradigm) would allow for an examination of these assumptions. At the moment, however, " […] the hypothesis that women show higher responses in interpersonal tasks remains speculative." (Kudielka & Kirschbaum, 2005, p. 125).

Lastly, sex differences in facial expression recognition should be taken into account in future research. Generally, studies support the hypothesis of a better recognition of emotional facial expressions in females compared to males (c.f., Kryspin-Exner & Felnhofer, 2012; McClure, 2000). The fact that facial expressions were limited in our study might explain the lack of sex differences in stress responses. Research that experimentally manipulates facial expressions of virtual others may provide additional evidence on sex specific attentional processes and vigilance-avoidance patterns.

5. Limitations

In sum, the current study only presents a particular type of social stimulus, a social-evaluative scenario with limited social interaction; in other words, there was no complex or verbal social interaction like in Hartanto et al. (2014). Thus, in future studies the number of scenarios should be increased to include different kinds of socially stressful situations (i.e. social-evaluative, social exclusion etc.). Given that individuals' reactions to social stimuli may differ depending on the type of stimulus used (c.f. Kampmann et al., 2016), presenting a range of situations, especially more complex situation allowing verbal interaction with virtual social entities, may provide a more customized approach to VRET and ensure its effective use.

6. Conclusion

Our results are particularly encouraging with regards to the use of VR applications in research and therapy. In neuroscience and stress research VR stimuli may be validly and reliably used to generate reactions in users that satisfy both the experimental rigor and ecological validity. Furthermore, given that the mere presence of virtual entities with limited facial expressions and interaction possibilities may already evoke strong social stress responses in users, VR is also valuable for therapeutic approaches. For instance, in cognitive behavior therapy (CBT) based exposure, it is crucial that the response to a stimulus is sufficiently strong to provide ground for processes such as extinction (c.f., Jönsson et al., 2010; Krijn et al., 2004). Finding strong responses on a multidimensional assessment of social stress (self-report, HPA- and SAM axis), our study demonstrates that virtual social stimuli may satisfy this precondition.

Another implication of the observed correspondence in stress reactivity to both a virtual and a real-life stimulus becomes apparent when reverting to MMORPGs and collaborative virtual environments (CVEs) such as Second Life. These platforms provide individuals with the unique possibility to engage in social interactions with a wide range virtual others; for instance, MMORPGs have previously been shown to promote social acceptance and social integration especially for lonely or shy individuals (Stetina, Kothgassner, Lehenbauer, & Kryspin-Exner, 2011). On the basis of the current results, however, it may be assumed that if a socially threatening situation is experienced in MMORPGs or in CVEs it may have a similar effect on an individual as a real-life social threat. Studies examining the effects of virtual social exclusion and ostracism support this notion (Kassner et al., 2012; Kothgassner et al., 2014). They show that a virtually ostracized person has comparable negative reactions to it as a person who is socially excluded in real life.

With respect to the technological developments of the past years, it can be assumed that immersive technology will play a major role in future communication and entertainment applications. When social network sites will migrate to virtual reality, it will profit from the rapidly growing enhancement of behavioral realism and the higher degrees of interactivity in online gaming what will continuously close the gap between the virtual reality and physical reality (Bailenson & Blascovich, 2011). This underlines the importance to understand social interactions in virtual reality in general more deeply, especially compared to interactions with physically present social entities in real life, to estimate the influence of virtual social stimuli on the experience of and behavior in a person's real life.

Acknowledgments

The authors of this work want to thank Jasmine Gomm and Mag. Anna Katharina Heinzle for helping with the data entry for the ECG recordings. Furthermore, we thank Mag. Janka Scharfenberger, Mag. Nora Pösl, and Mag. Anna Hoffmann for their assistance in the acquisition of some parts of the data and Edith Klobetz-Rassam for cortisol analysis. Also we would like to thank Dr. Mario Lehenbauer for his support in the early stages of the project.

References

Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E., et al. (2013). Virtual reality exposure therapy for social anxiety disorder: a randomized controlled trial. *Journal of Consulting and Clinical Psychology, 81*(5), 751–760.

Bailenson, J. N., & Blascovich, J. (2011). Virtual reality and social networks will be a powerful combination: avatars will make social networks seductive. *IEEE Spectrum*. http://spectrum.ieee.org/telecom/internet/virtual-reality-and-social-networks-will-be-a-powerful-combination.

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin, 29*, 819–833.

Bente, G., Krämer, N. C., & Petersen, A. (Eds.). (2002). *Virtuelle Realitäten*. Hogrefe Verlag.

Bigger, J. T., Fleiss, J. L., Steinman, R. C., Rolnitzky, L. M., Schneider, W. J., & Stein, P. K. (1995). RR variability in healthy, middle-aged persons compared with patients with chronic coronary heart disease or recent acute myocardial infarction. *Circulation, 91*, 1936–1943.

Bitsika, V., Sharpley, C. F., Sweeney, J. A., & McFarlane, J. R. (2014). HPA and SAM axis responses as correlates of self-vs parental ratings of anxiety in boys with an autistic disorder. *Physiology & Behavior, 127*, 1–7.

Bordnick, P. S., Traylor, A. C., Carter, B. L., & Graap, K. M. (2012). A feasibility study of virtual reality-based coping skills training for nicotine dependence. *Research on Social Work Practice, 22*, 293–300.

Brom, C., Buchtová, M., Šisler, V., Děchtěrenko, F., Palme, R., & Glenk, L. M. (2014). Flow, social-interaction anxiety and salivary cortisol responses in serious games: a quasi-experimental study. *Computers & Education, 79*, 69–100.

Chen, Y. P., Ehlers, A., Clark, D. M., & Mansell, W. (2002). Patients with generalized social phobia direct their attention away from faces. *Behaviour Research and Therapy, 40*(6), 677–687.

Cummings, J. J., & Bailenson, J. N. (2015). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, 1–38.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*(3), 355–389.

Dishman, R. K., Nakamura, Y., Garcia, M. E., Thompson, R. W., Dunn, A. L., & Blair, S. N. (2000). Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. *International Journal of Psychophysiology, 37*(2), 121–133.

Felnhofer, A., Kothgassner, O. D., Beutl, L., Hlavacs, H., & Kryspin-Exner, I. (2012). Is virtual reality made for men only? Exploring gender differences in the sense of presence. In *Proceedings of the International Society on presence research* (Philadelphia, USA).

Felnhofer, A., Kothgassner, O. D., Hetterle, T., Beutl, L., Hlavacs, H., & Kryspin-Exner, I. (2014). Afraid to be there? Evaluating the relation between presence, self-reported anxiety, and heart rate in a virtual public speaking task. *Cyberpsychology, Behavior, and Social Networking, 17*(5), 310–316.

Felnhofer, A., Kothgassner, O. D., Schmidt, M., Heinzle, A. K., Beutl, L., Hlavacs, H., et al. (2015). Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. *International Journal of Human-Computer Studies, 82*, 48–56.

Garau, M., Slater, M., Pertaub, D.-P., & Razzaque, S. (2005). The responses of people to virtual humans in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments, 14*, 104–116. http://dx.doi.org/10.1162/1054746053890242.

Garcia-Lopez, L., Olivares, J., Beidel, D., Albano, A., Turner, S., & Rosa, A. (2006). Efficacy of three treatmetn protocols for adolescents with social anxiety disorder: a 5-year follow-up assessment. *Journal of Anxiety Disorders, 20*, 175–191.

Garcia-Palacios, A., Botella, C., Hoffman, H., & Fabregat, S. (2007). Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. *Cyberpsychology & Behavior, 10*(5), 722–724.

Gerardi, M., Cukor, J., Difede, J., Rizzo, A., & Rothbaum, B. O. (2010). Virtual reality exposure therapy for post-traumatic stress disorder and other anxiety disorders. *Current Psychiatry Reports, 12*(4), 298–305.

Girden, E. (1992). *ANOVA: Repeated measures*. Newbury Park, CA: Sage.

Glaser, R., & Kiecolt-Glaser, J. K. (2005). Stress-induced immune dysfunction: implications for health. *Nature Reviews Immunology, 5*(3), 243–251.

Harris, S. R., Kemmerling, R. L., & North, M. M. (2002). Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology & Behavior, 5*(6), 543–550.

Hartanto, D., Kampmann, I. L., Morina, N., Emmelkamp, P. G., Neerincx, M. A., & Brinkman, W. P. (2014). Controlling social stress in virtual reality environments. *PLoS One, 9*(3), e92804.

Hemmeter, U., Störmer, R., Mager, R., Kuntze, M., Mueller-Spahn, F., Hennig, J., et al. (2005). Modification of virtual reality combined with a mental task stimulates cortisol in volunteers. *Neuropsychobiology, 51*, 165–172.

Hofmann, S. G. (2007). Cognitive factors that maintain social anxiety disorder: a comprehensive model and its treatment implications. *Cognitive Behaviour Therapy, 36*(4), 193–209.

Jönsson, P., Wallergård, M., Österberg, K., Hansen, Å. M., Johansson, G., & Karlson, B. (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.

Kampmann, I. L., Emmelkamp, P. M., Hartanto, D., Brinkman, W. P., Zijlstra, B. J., & Morina, N. (2016). Exposure to virtual social interactions in the treatment of social anxiety disorder: a randomized controlled trial. *Behaviour Research and Therapy, 77*, 147–156.

Kang, N., Brinkman, W. P., van Riemsdijk, M. B., & Neerincx, M. (2016). The design of virtual audiences: noticeable and recognizable behavioral styles. *Computers in Human Behavior, 55*, 680–694.

Kassner, M. P., Wesselmann, E. D., Law, A. T., & Williams, K. D. (2012). Virtually ostracized: studying ostracism in immersive virtual environments. *Cyberpsychology, Behavior, and Social Networking, 15*(8), 399–403.

Kelly, O., Matheson, K., Martinez, A., Merali, Z., & Anisman, H. (2007). Psychosocial stress evoked by a virtual audience: relation to neuroendocrine activity. *CyberPsychology & Behavior, 10*(5), 655–662.

Kelly, M. M., Tyrka, A. R., Anderson, G. M., Price, L. H., & Carpenter, L. L. (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.

Kemp, A. H., Quintana, D. S., Felmingham, K. L., Matthews, S., & Jelinek, H. F. (2012). Depression, comorbid anxiety disorders, and heart rate variability in physically healthy, unmedicated patients: implications for cardiovascular risk. *PLoS One, 7*(2), e30777.

Kirschbaum, C., Kudielka, B. M., Gaab, J., Schommer, N. C., & Hellhammer, D. H. (1999). Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic Medicine, 61*(2), 154–162.

Klinger, E., Bouchard, S., Légeron, P., Roy, S., Lauer, F., Chemin, I., et al. (2005). Virtual reality therapy versus cognitive behavior therapy for social phobia: a preliminary controlled study. *Cyberpsychology & Behavior, 8*(1), 76–88.

Kothgassner, O. D., Kafka, J., Rudyk, J., Beutl, L., Hlavacs, H., & Felnhofer, A. (2014). Does social exclusion hurt virtually like it hurts in real life? The role of agency and social presence in the perception and experience of social exclusion. *Proceedings of the International Society for Presence Research*, 45–56.

Kotlyar, M., Donahue, C., Thuras, P., Kushner, M. G., O'Gorman, N., Smith, E. A., et al. (2008). Physiological response to a speech stressor presented in a virtual reality environment. *Psychophysiology, 45*(6), 1034–1037.

Krijn, M., Emmelkamp, P. M., Olafsson, R. P., & Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: a review. *Clinical Psychology Review, 24*(3), 259–281.

Kryspin-Exner, I., & Felnhofer, A. (2012). Zur Psychologie des "kleinen Unterschieds". Betrachtung von Gender-Merkmalen über die Lebensspanne [The Psychology of a „little difference". Gender aspects over the life-span]. In A. Kautzky-Willer (Ed.), *Gender Medizin: Personalisierte Medizin für Mann und Frau [Gender medicine: Personalized medicine for men and women]* (pp. 17–32) (Wien Köln Weimar: Böhlau UTB).

Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H., & Kirschbaum, C. (2004). HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: impact of age and gender. *Psychoneuroendocrinology, 29*(1), 83–98.

Kudielka, B. M., Hellhammer, D. H., & Wüst, S. (2009). Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology, 34*(1), 2–18.

Kudielka, B. M., & Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biological Psychology, 69*(1), 113–132.

Kudielka, B. M., Schommer, N. C., Hellhammer, D. H., & Kirschbaum, C. (2004). Acute HPA axis responses, heart rate, and mood changes to psychosocial stress (TSST) in humans at different times of day. *Psychoneuroendocrinology, 29*(8), 983–992.

Lehenbauer, M., Kothgassner, O. D., Kryspin-Exner, I., & Stetina, B. U. (2013). An online self-administered social skills training for young adults: results from a pilot study. *Computers & Education, 61*, 217–224.

Mansell, W., Clark, D. M., Ehlers, A., & Chen, Y. P. (1999). Social anxiety and attention away from emotional faces. *Cognition & Emotion, 13*(6), 673–690.

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin, 126*(3), 424.

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine, 338*, 171–179.

McEwen, B. S. (2005). Glucocorticoids, depression, and mood disorders: structural remodeling in the brain. *Metabolism, 54*(5), 20–23.

Montero-López, E., Santos-Ruiz, A., García-Ríos, M. C., Rodríguez-Blázquez, R., Pérez-García, M., & Peralta-Ramírez, M. I. (2015). A virtual reality approach to the trier social stress test: contrasting two distinct protocols. *Behavior Research Methods*, 1–10.

Morina, N., Ijntema, H., Meyerbröker, K., & Emmelkamp, P. M. G. (2015). Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behaviour Research and Therapy, 74*, 18–24.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: social responses to computers. *Journal of Social Issues, 56*(1), 81–103.

Opris, D., Pintea, S., Garcia-Palacios, A., Botella, C., Szamosk, A., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: a quantitative meta-analysis. *Depression & Anxiety (1091-4269), 29*, 85–93. http://dx.doi.org/10.1002/da.2091.

Owens, M. E., & Beidel, D. C. (2015). Can virtual reality effectively elicit distress associated with social anxiety disorder? *Journal of Psychopathology and Behavioral Assessment, 37*(2), 296–305.

Palme, R., & Möstl, E. (1997). Measurement of cortisol metabolites in feces of sheep as a parameter of cortisol concentration in blood. *International Journal of Mammal Biology, 62*(Suppl. 2), 192–197.

Parsons, T. D., & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(3), 250–261.

Paul, G. (1966). *Insight vs. desensitization in psychotherapy*. Stanford University Press.

Pertaub, D. P., Slater, M., & Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators and Virtual Environments, 11*(1), 68–78.

Pittig, A., Arch, J. J., Lam, C. W., & Craske, M. G. (2013). Heart rate and heart rate variability in panic, social anxiety, obsessive–compulsive, and generalized anxiety disorders at baseline and in response to relaxation and hyperventilation. *International Journal of Psychophysiology, 87*(1), 19–27.

Powers, M. B., & Emmelkamp, P. M. (2008). Virtual reality exposure therapy for anxiety disorders: a meta-analysis. *Journal of Anxiety Disorders, 22*(3), 561–569.

Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology, 28*(7), 916–931.

Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.

Repetto, C., Gaggioli, A., Pallavicini, F., Cipresso, P., Raspelli, S., & Riva, G. (2011). Virtual reality and mobile phones in treatment of generalized anxiety disorders: a phase-2 clinical trial. *Personal and Ubiquitous Computing, 17*, 253–260.

Sarver, N. W., Beidel, D. C., & Spitalnick, J. S. (2014). The feasibility and acceptability of virtual environments in the treatmetn of childhood social anxiety disorder. *Journal of Clinical Child & Adolescent Psychology, 43*, 63–73. http://dx.doi.org/10.1080/15374416.2013, 843461.

Schommer, N. C., Hellhammer, D. H., & Kirschbaum, C. (2003). Dissociation between reactivity of the hypothalamus-pituitary-adrenal axis and the sympathetic-adrenal-medullary system to repeated psychosocial stress. *Psychosomatic Medicine, 65*(3), 450–460.

Slater, M., Pertaub, D. P., Barker, C., & Clark, D. M. (2006). An experimental study on fear of public speaking using a virtual environment. *CyberPsychology & Behavior, 9*(5), 627–633.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *State-trait anxiety inventory (self-evaluation questionnaire)*. Consulting Psychololgists Press.

Stetina, B. U., Kothgassner, O. D., Lehenbauer, M., & Kryspin-Exner, I. (2011). Beyond the fascination of online-games: probing addictive behavior and depression in the world of online-gaming. *Computers in Human Behavior, 27*(1), 473–479.

Stroud, L. R., Salovey, P., & Epel, E. S. (2002). Sex differences in stress responses: social rejection versus achievement stress. *Biological Psychiatry, 52*(4), 318–327.

Szell, M., & Thurner, S. (2013). How women organize social networks different from men. *Scientific Reports, 3*, 1214.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Circulation, 93*, 1043–1065.

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. *Psychological Review, 107*(3), 411.

Villani, D., Repetto, C., Cipresso, P., & Riva, G. (2012). May I experience more presence in doing the same thing in virtual reality than in reality? An answer from a simulated job interview. *Interacting with Computers, 24*(4), 265–272.

Von der Pütten, A., Krämer, N. C., Gratch, J., & Kang, S. H. (2010). "It doesn't matter what you are!" Explaining social effects of agents and avatars. *Computers in Human Behavior, 26*(6), 1641–1650.

Williams, K. D., & Jarvis, B. (2006). Cyberball: a program for use in research on interpersonal ostracism and acceptance. *Behavior Research Methods, 38*(1), 174–180.

Wolpe, J. (1961). The systematic desensitization treatment of neuroses. *The Journal of Nervous and Mental Disease, 132*(3), 189–203.

Yuen, E. K., Herbert, J. D., Forman, E. M., Goetter, E. M., Comer, R., & Bradley, J. C. (2013). Treatment of social anxiety disorder using online virtual environments in second life. *Behavior Therapy, 44*(1), 51–61.