Examination of electrodermal and cardio-vascular reactivity in virtual reality through a combined stress induction protocol

Bojan Kerous¹ · Richard Barteček¹ · Robert Roman¹ · Petr Sojka¹ · Ondřej Bečev¹ · Fotis Liarokapis¹

Received: 14 October 2019 / Accepted: 28 February 2020 / Published online: 9 March 2020
© The Author(s) 2020

Abstract

In this study, task-related stress induction through Stroop task and social stress induction protocol based on elements of Trier Social Stress Test are examined. The aim of the paper is to find the optimal combination of social and task-related stress to be used to consistently and reliably induce a stressful reaction. In total 16 healthy subjects participated in this study that seeks to find and compare the different stressors and their relation to physiological reactivity. Our findings show that electrodermal activity measurements are suitable when using a combination of stressors while heart rate and Root Mean Square of the Successive Differences highlight a greater reactivity to task-stress.

Keywords Virtual reality · Social stress · Anxiety disorders

1 Introduction

In the real-world stress occurs whenever a demand exceeds the regulatory capacity of an organism, particularly in situations that are unpredictable and uncontrollable. In experimental stress induction procedures these stressful conditions are emulated in a controlled and safe environment that typically contains elements of some physical challenge, a cognitively demanding tasks or a social-evaluative threat (Starcke and Brand 2012). A disrupted reactivity to stressful stimuli as seen in anxiety disorders is reflected in the deregulation of the functioning of the Autonomic Nervous System (ANS) whose task is to maintain appropriate reactivity of bodily functions to outside stressors. For the extinction of the pathological stressful response to occur the patient typically undergoes behavioral therapy that includes controlled exposure to the stressor and is gradually trained in recognizing and managing the effects. Considering that even a neutral stimuli can cause a stress reaction and cause pathological responses for decades after danger has passed (VanElzakker et al. 2014) and the fact that behavioral therapy itself can lead to worsening of the symptoms (Lupien et al. 2018), there needs to be considerable care in designing

the virtual environments, as well as in the choice of stressors and corresponding physiological measures.

While virtual reality (VR) provides many benefits, our knowledge of the language of the medium and the tools that are available is not sufficient to guaranty validity or efficacy of a specific application. Therefore, exploratory studies that examine different stressor sources and modalities of assessment of their utility are merited. Given that higher presence is reported when emotional stimulus is used (Diemer et al. 2015) in the case of VR-mediated behavioral therapy the totality of the stimulus space has to consistently evoke intended reactions through repeated exposures, while avoiding situations that would induce seizures or simulation sickness and thereby disrupt the intervention. Furthermore, technological disruption due to e.g. loss of head tracking or synchronization of events leads to deterioration of the effectiveness of stress induction (Pallavicini et al. 2013) and significant advantages offered by VR disappear and its emotional induction abilities are even lower than the ones provided by cheaper media.

Even though the field of VR-based research geared toward clinical application is not a novelty, there is still no direct correspondence of presence and related factors to the level of induced emotional arousal. Although a reinforcing role of realism on the behavioral response is noted in Slater et al. (2009) and in full body avatars in Gutiérrez-Maldonado et al. (2009), in Volonte et al. (2016) an opposite effect is reported. Furthermore, compared to alternatives, the

✉ Fotis Liarokapis
liarokap@fi.muni.cz

¹ Masaryk University, Brno, Czech Republic

effectiveness of head-mounted display (HMD)-based exposure is not clear as reported in Calogiuri et al. (2018) and Kim et al. (2014). Given the enormity of the leap from the omnipresent desktop interaction techniques to an unexplored language of interaction in VR, the design process associated with such media is expected to evolve accordingly. This has been a topic of debate in Blom and Beckhaus (2014) where a design taxonomy has been put forward and in Sutcliffe et al. (2019) where VR design process is examined in the context of standard Graphical User Interface (GUI). Despite the breadth of the research there are no clear recommendations towards a one-size-fits-all intervention.

While VR has proven advantages over conventional means, our knowledge of the language of the medium and the tools that are available is not sufficient to guaranty validity or efficacy of a specific application. Therefore, exploratory studies that examine different stressor sources and modalities of assessment of their utility are needed. In order to make informed decisions on the design of our stress induction procedure we did not rely on any stringent specification of similar interventions as no such methodological guidelines exist. Instead, we have embarked on an iterative approach using smaller groups of healthy participants. Throughout the design phase the development was overseen by a combined team of VR experts and practicing therapists with the priority of establishing a safe and reliable stress induction protocol.

This paper aims at establishing a reliable stress induction protocol and examining physiological reactivity in a nuanced and controlled experimental setup. This involves a moderate social stressor, a control stressful condition (Stroop test in VR), and a combination of both conditions, with initial physiological measurements taken from a non-threatening virtual park environment. Our findings indicate that electrodermal activity (EDA) measurements, namely the number of non-specific skin conductance responses (nSCR), better indicate stress reaction when a combination of stressors is used, while heart rate (HR) and Root Mean Square of the Successive Differences (RMSSD) better react to task-stress.

The rest of the paper is structured as follows. Section 2 presents the motivation of the study where Sect. 3 the methodology followed. Section 4 presents our experimental results. Section 5, provides a discussion of the results and Sect. 6 the limitations. Finally, Sect. 7 presents the conclusions.

2 Motivation

The basis of our approach is on using two stress induction procedures to, firstly assess their effectiveness in stress induction, and secondly, to contrast and combine their effects to achieve reliable stress induction capability. Ever

since very early experiments with anxiety in VR, induction of social stress through virtual character's behaviors has been a promising prospect. As authors note in an early work on the topic (Slater et al. 2006) the stressful social environment represents relatively normal, everyday conditions as opposed to more violent stressors such as a wartime environment. For phobic people this apparently ordinary situation is enough to induce a stress response and may therefore be a more ecologically valid means of behavioral conditioning.

Nowadays, a common method of inducing social stress is the Trier Social Stress Test (TSST) (Kirschbaum et al. 1993). The protocol, originally conducted with live actors, consisted of a job interview, and was followed by a mental arithmetic task. Given its utility (Kelly et al. 2007) and its consistency in inducing physiological responses (Ruiz et al. 2010), the protocol has seen numerous customizations to meet the requirements of stress induction in specific cases (Allen et al. 2017). TSST protocol also has several limitations that may extend to its VR counterpart (Frisch et al. 2015). The scope of the method is limited to social-evaluative threat, features an additional mental arithmetic task which makes it difficult to evaluate independent effects, and results in habituation to stressor.

To address these limitations and compare the magnitude of the elicited social stress it was important to introduce the appropriate benchmark task. For this purpose, the Stroop task (Stroop 1935) was considered as it is a widely used task with a potential to indicate a variety of cognitive processes and has the potential to consistently and repeatedly induce stress with limited or no effect of habituation. The Stroop is a color naming task where a person is shown written color names, where the color of the text can be in accordance with the text which makes identifying the color easy, or it can be incongruent with it. In the original experiment five words (red, blue, green, brown and purple) and five corresponding colors are used (no instance does the word match the color of the text). It is held that this effect has good resilience to habituation which makes possible repeated exposure with reliable outcomes.

VR Stroop task was first reported in (Rizzo et al. 2009) where it was embedded in a virtual classroom environment. In a drastically different example, it was added to a warzone simulation in Wu et al. (2010) where performance scores in the test were in accord with traditional tests of executive functioning (Armstrong et al. 2013). The addition of the Stroop task not only provides insight into task performance but also augments the environmental and other cue in inducing the desired response (Parsons and Reinebold 2012). For the purposes of this research the Stroop task was used as a replacement of the mental arithmetic task that is typically present in TSST intervention, with the social stressors being reduced to quiet scrutiny. To remove motion sickness

effects, this work was based on concepts presented at Luks and Liarokapis (2019).

As was reported in Norrholm et al. (2016) establishing a physiological reactivity for a specific set of virtual scenes can help guide subsequent cognitive behavioral intervention, as the outcome of such therapies can then be assessed through the reduced physiological reactivity post-treatment. To assess the physiological reactivity two psycho-physiological measurements were used to assess the functioning of the ANS. Heart rate variability (HRV) measures are a widely utilized indicators of attentional regulation (Thayer and Lane 2009) and affective information processing and dis-regulation measured through HRV is reported in both patients with Generalized Anxiety disorder and in healthy people and patients during worry (Thayer and Lane 2000).

On the other hand, as clarified by functional imaging research, the brain regions implicated in emotion, attention, and cognition play a role in EDA responses (Critchley 2002). Specifically a continuous measure of phasic activity was proposed in Benedek and Kaernbach (2010) as a straight-forward indicator of event-related sympathetic activity. Taken together EDA and ECG have before been used to assess quality of experience (QoE) in VR content and are generally

inexpensive and unobtrusive in a VR intervention compared to alternatives (Egan et al. 2016).

3 Methodology

Testing was carried in three steps which featured the same environments (illustrated in Fig. 1) with the only distinction being the ordering of the scenes. In step 2, two different orders of VR scenes were used, in further text, these variants are considered as separate experiments 2A and 2B. The overview of the testing is shown in Table 1, which presents a summary of several subjects enrolled in each experiment of the study as well as an order of different VR scenes in specific experiments. Numbers noted under specific scene names indicate the order of scenes while “NA” indicates that the scene was not used during the experiment. “Park1” and “Park2” conditions represents the first and second appearance in the park scene, the “Social” condition marks the scene where only social stressors (virtual characters are present in the hospital office scene), the “Stroop” condition stands for the same scene without characters, and with the Stroop task, while “Social Stroop” represents a combination

Fig. 1 Prior to initializing the experiment, the experimenter can customize duration and ordering of subsequent scenes, as well as the disclaimer text shown to participants between scenes (Image 1). Subjects were first instructed to relax in a virtual park scene (Image 2) where they got acquainted with the visualization of their hands as well as with basic interaction. Upon the completion of the allotted time spent

in the park, and once the pop-up disclaimer is read, the participant can acknowledge the information and proceed to the next scene by pressing the virtual button (Image 3). The following scene was the hospital environment (Image 4) where the subjects were presented with the Stroop task and exposed to the socially stressful virtual character behavior

Table 1 Overview of the experiments

Experiment	Subjects	Order of scenes				
		Park 1	Social	Stroop	Social Stroop	Park 2
1	7	1	NA	2	3	4
2A	5	1	3	NA	2	4
2B	5	1	2	NA	3	4

of social and task stressors. The hypotheses of this work are as follows:

- H1: Our first hypothesis for this phase of the study was that there would be statistically significant differences in physiological measurements in different VR scenes.
- H2: The second hypothesis was that the VR scenes used in the experiment and their order (in other words experiment type) do not influence stress reaction.

In total, 17 healthy subjects were enrolled into these experiments (see Table 1). However, one subject was excluded due to probable mistakes in measurements, because in this subject CDA nSCR, HR and RMSSD had zero values in all scenes. The total number of subjects entering subsequent analyses is therefore 16. All four physiological measurements were used in this phase of the study: count of non-specific responses in EDA (CDA nSCR), tonic EDA (CDA Tonic), mean HR and RMSSD.

The initial scene was a park scene which was used to obtain baseline values for electro-dermal and cardiovascular measures while the participants are instructed to relax. Subject's hands were tracked using a Leap Motion infrared sensor (<http://www.leapmotion.com>) mounted to the front of the HTC Vive Head-Mounted Display (HMD) (<http://www.vive.com>). The users were initiating scene changes themselves by interacting with the disclaimers at the beginning and end of every scene. The disclaimers displayed simple instructions in Czech language introducing them to basic interaction, as well as the task in the scene ahead. The disclaimer is removed by pressing a button (with the virtual hand) which in turn initiated the timer for the scene.

All the virtual characters were created to represent unspecified medical workers on various levels of job hierarchy. In Experiments 1 and 2, the characters remain quiet and seated close to the participant while they maintain eye contact (as reported in Krejsa et al. 2018). To achieve a satisfactory visual fidelity and enable quick prototyping of virtual scenarios, the virtual characters were generated using the commercial package Adobe Fuse (<http://www.adobe.com/products/fuse.html>) and animated using Mixamo library (<http://www.mixamo.com/>) of character animations.

Four types of measurements were selected to be analyzed: count of non-specific responses in EDA (CDA nSCR),

tonic EDA (CDA Tonic), mean heart rate (HR), RMSSD as a measurement of heart-rate variability. These measurements were obtained for each VR scene. EDA and electrical activity of the heart (ECG) were recorded with Psychlab data acquisition system (<http://www.psychlab.com>) with a sampling rate of 1000 Hz. Measurement was performed in a quiet shielded room with a temperature of 23 °C.

For EDA recording, one pair of Ag/AgCl electrodes with 8 mm active area diameter filled with electroconductive paste was used. The electrodes were attached to the medial phalanges of the middle and index finger of the left hand. ECG electrodes were positioned on the chest at convenient sides of the heart. Ledalab software was used for a decomposition of EDA signal into its tonic and phasic components according to continuous decomposition analysis (Benedek and Kaernbach 2010). HRV of ECG data were analysed with Kubios HRV Premium software (Tarvainen et al. 2014).

4 Results

Results of physiological measurements were used as dependent variables while independent variables differed between experiments. In all experiments, all four physiological measurements were tested (CDA nSCR, CDA Tonic, HR, RMSSD). Distribution of physiological measurements was tested by the Shapiro–Wilk normality test at the p-value of 0.05. Because the distribution of neither variable was normal, physiological measurements were rank-transformed before further analyses. The subsequent analysis had several steps: The first step was selecting optimal models: Independent variables were gradually introduced into models thus creating models with gradually increasing complexity.

In total four groups of models were explored, corresponding with four dependent variables. Models were compared according to Akaike information criterion (AIC) and Likelihood ratio. Smaller AIC was considered as a measurement of a better model. Differences between models were described by Likelihood ratio and were statistically tested. Model with a lowest AIC and statistically significant difference when comparing to the model of previous complexity were selected as models entering further analyses. In further analyses, Bonferroni correction was used with a correction factor of four reflecting the number of physiological variables explored. The p-value for further tests was therefore

lowered to 0.0125 ($0.05/4 = 0.0125$). Both, results with and without Bonferroni correction were reported. The analysis was performed using R statistical library with the “Nlme” package for mixed effects models and the “Multcomp” package for post hoc test (multiple comparisons of means—Tukey contrasts).

In mixed effects models, VR scene (Scene) and experiment (Experiment) were selected as possible predicting fixed effects variables while subjects were considered a random effect variables. The Scene variable was used to assess the capacity of different VR scenes to elicit stress reaction (H1), the Experiment variable was used to evaluate the effect of a different set of VR scenes and their order (H2). In all physiological measures, the optimal models contained Scene as a dependent variable. However, in the case of CDA Tonic, the model with both Scene and Experiment appeared to explain the physiological measurement response better. The result of the model selection is depicted in Table 2. In the case of CDA nSCR, HR and RMSSD the Scene statistically significantly explained their changes. However, both Scene and Experiment statistically significantly predicted CDA Tonic. The results of the mixed effects models are summarized in Table 3.

In CDA nSCR the Stroop and Social Stroop scenes differed statistically significantly from both park scenes (Park 1 and 2). Social scene differed significantly only when compared to Park 2. In CDA Tonic Park 1, Social Stroop and Park 2 differed significantly from each other. Social and Stroop scene did not differ from other scenes. In HR, Stroop and Social Stroop showed statistically significantly higher values than all other scenes. The result was similar in RMSSD, however here the Stroop and Social Stroop showed significantly lower values from all other scenes. The results

Table 3 Mixed effects models results

Dependent variable	Scene		Experiment	
	F-value	P-value	F-value	P-value
CDA nSCR	17.38	< 0.001***		
CDA nSCR	16.79	< 0.001***	5.42	0.019
CDA Tonic	11.21	< 0.001***		
CDA Tonic	10.84	< 0.001***	18.62	< 0.001***
HR	34.19	< 0.001***		
HR	33.03	< 0.001***	0.61	0.561
RMSSD	21.01	< 0.001***		
RMSSD	20.28	< 0.001***	0.80	0.469

Statistically significant differences are highlighted: ***The difference is statistically significant at $p < 0.0125$ (Bonferroni corrected)

of post-hoc tests are summarized in Table 4 and Fig. 2, the values of physiological measures in different VR scenes are depicted in Table 5 and Fig. 2.

5 Discussion

Research on VR is accelerating rapidly across applications. Illustrating the scope, in a 2018 paper (Cipresso et al. 2018) authors have gathered 21667 publications relating to VR within the Web of Science Core collection alone. Authors note an increase in journal publications on the topic as well as an increase in clinical inquiries. Amid such a burgeoning field it will become more difficult to generalize outcomes especially with continuous and accelerated evolution of VR interfaces. As technology-driven research is not sensitized to the patient's needs, there is a necessity to put greater

Table 2 Comparison of different models

Model	DV	IV	AIC	Models compared	L-ratio	P-value
A1	CDA nSCR		604.45	NA	NA	NA
A2†	CDA nSCR	Sc	598.09	A1–A2	14.36	0.006***
A3	CDA nSCR	Sc	601.09	A2–A3	1.00	0.605
B1	CDA Tonic		502.71	NA	NA	NA
B2	CDA Tonic	Sc	476.43	B1–B2	34.29	< 0.001***
B3†	CDA Tonic	Sc	457.30	B2–B3	23.12	< 0.001***
C1	HR		582.39	NA	NA	NA
C2†	HR	Sc	528.01	C1–C2	62.39	< 0.001***
C3	HR	Sc	531.61	C2–C3	0.40	0.819
D1	RMSSD		583.49	NA	NA	NA
D2†	RMSSD	Sc	544.18	D1–D2	47.31	< 0.001***
D3	RMSSD	Sc	545.32	D2–D3	2.86	0.239

Statistically significant differences are highlighted: ***The difference is statistically significant at $p < 0.0125$ (Bonferroni corrected). The best model is highlighted with †. DV dependent variable, IV independent variable, AIC akaike information criterion, L-ratio Likelihood ratio, CDA nSCR electrodermal activity-number of non-specific responses, CDA Tonic electrodermal activity-tonic, HR heart rate, RMSSD Root mean square of successive differences, Sc scene, Exp experiment, NA not applicable

Fig. 2 Physiological variables during different scenes. The predictors used in final models are in brackets. Letters above each box represent different groups according to a statistical difference. Groups were assigned according to post hoc multiple comparisons of means (Tukey contrasts) at $p < 0.0125$ (Bonferroni corrected)

emphasis on user-centered design (Kellmeyer 2018), especially due to the fact that most VR systems nowadays are geared towards entertainment applications and may be limited in several aspects from the requirements of clinical use.

Use of VR in behavioral therapy is still in early stages partly due to the fact that therapists may still view the technology as immature. However, after the 2016 release of several HMD devices, clinicians seem to have less reservations in using commercially available VR hardware, as reported in Lindner et al. (2019). Authors show showed that attitudes of polled therapist's ($n = 185$) toward VR in exposure therapy appear to have evolved in recent years. Although, HMDs of suitable specifications may be available already greater flexibility is required in the software delivered to the lab (Cipresso et al. 2018) with respect to offering multiple virtual environments, stressors or interaction modalities. In this respect our work may prove to be a valuable example of such flexibility which was achieved by developing a protocol which can be customized without requiring the clinical staff to have specific technical knowledge.

The effectiveness of VR stress induction in therapeutics has been the topic of a substantial number of meta-analyses through the years each adding weight to the argument. In fact, the bulk of its reported use was for the purpose of treating a wide array of anxiety disorders (Maples-Keller et al. 2017). This is substantiated also in Carl et al. (2019), a meta analysis of randomized controlled trials encompassing 30 studies in total (participants $N=1057$) that reported a large

effect size compared to waitlist controls for VR based therapy and had found that VR exposure is no less effective than in vivo exposure, with the consistent effect sizes across different affective disorders. In another meta-analysis (Benbow and Anderson 2019) (46 studies, $N=1057$) the attrition rates of such treatment was examined, with the conclusion that with the attrition rate of 16% it is brought in line with the attrition rates seen in in-vivo treatments. Even though there is ample evidence of usefulness of VR induction, common stress induction techniques used in traditional behavioral therapy are subject to specific restrictions and merit caution due to the immersive nature of the interface.

This is even more true for clinical applications where treatment outcomes directly depend on the design process. Among the stakeholders such as therapists there are reported reservations towards technology adoption (Schwartzman et al. 2012). The feedback suggests the therapists are reserved due to the requirement of training, special equipment, cost and general unfamiliarity with benefits VR provides. Nevertheless, VR therapists hold a positive attitude towards the technology with high cost not being an issue, and greater emphasis put on translatability to real-life outcomes (Lindner et al. 2019).

For a VR approximation of a real-life situation to be effective, several conditions need to be met. In Slater (2009), authors state these conditions to be reflected in the reported notion of "being there" (known as presence), as well as the perceived plausibility of the environment, while

Table 4 Results of post hoc multiple comparisons of means (Tukey contrasts) on final models

	Park 1	Social	Stroop	Social Stroop	Park 2
CDA nSCR (scene)					
Park 1	1	0.01	0.01	< 0.001	0.59
Social	0.01	1	0.94	0.22	< 0.001
Stroop	0.01	0.94	1	0.91	< 0.001
Social Stroop	< 0.001	0.22	0.91	1	< 0.001
Park 2	0.59	< 0.001	< 0.001	< 0.001	1
CDA tonic (scene, experiment)					
Park 1	1	0.02	0.02	< 0.001	0.01
Social	0.02	1	0.98	0.04	1.00
Stroop	0.02	0.98	1	0.42	0.92
Social Stroop	< 0.001	0.04	0.42	1	0.00
Park 2	0.01	1.00	0.92	0.00	1
HR (scene)					
Park 1	1	1	< 0.001	< 0.001	0.83
Social	1	1	< 0.001	< 0.001	0.95
Stroop	< 0.001	< 0.001	1	1.00	< 0.001
Social Stroop	< 0.001	< 0.001	1.00	1	< 0.001
Park 2	0.83	0.95	< 0.001	< 0.001	1
RMSSD (scene)					
Park 1	1	0.99	< 0.001	< 0.001	0.99
Social	0.99	1	< 0.001	< 0.001	0.92
Stroop	< 0.001	< 0.001	1	0.98	< 0.001
Social Stroop	< 0.001	< 0.001	0.98	1	< 0.001
Park 2	0.99	0.92	< 0.001	< 0.001	1

The predicting variables used in final models are in brackets. Each cell shows a p-value obtained by the post hoc test of the differences in physiological variables between a selected pair of scenes

in Price et al. (2011) an additional factor of realness is also acknowledged. Even though the extent of inquiries into this problem are increasing, concerns raised in Rizzo et al. (1998) remain valid.

Of special relevance is the issue of not only measuring presence as a construct originating from VR domain, but prescribing a specific amount for an application in a clinical context. Furthermore, breaks in so-called Place Illusion (the sense of being in a virtual space) and Plausibility Illusion (the overall plausibility of the events, surroundings or experiences) are a proven cause of degradation of the experience which will have a detrimental effect resulting in noticeable changes in physiological indicators or user state (Skarbez et al. 2018a). In a review on the construct of presence and other related concepts in Skarbez et al. (2018b) authors note three ways of measuring it, either through administration of self-report presence questionnaires, behavioral and

physiological measures. While questionnaires are the most widely used means of measuring presence the relation of self-reported presence to specific physiological variables is inconclusive.

Given the distinct characteristics of specific disorders of affect it is not clear to what extent side-effects (cybersickness and aftereffects) will have an impact on the population. Treatment for a wide array of specific disorders of affect may yet prove to have different mechanisms and associated outcomes, which can be expected given that the triggering stimuli can range from a spider, an open space, or a common social situation. The characteristics of specific disorder dictate both the structure of the intervention as well as the visual/audio assets used in creating a treatment platform.

The focus of this research was to comparatively assess the scope of the VR stress induction through a limited but focused apparatus of commonly used and easily acquired physiological measures (ECG and EDA). The intensity of the social stress condition was limited to the continuous observation by the surrounding virtual characters for two main reasons. One being the gradual, exploratory nature of our inquiry where consistency, ecological validity and safety of the stress induction was a priority, and the other where the goal of contrasting and mixing stressors to achieve such ends is utilized with the augmentation through addition of a consistent task-stressor.

6 Limitations

The initial park scene was not chosen as a neutral condition, but served as generally passive environment that would promote relaxation in order to accommodate the participant to wearing the HMD and the visualization of their hands. Although the participants did not report any physical or visual discomfort when using the Leap Motion-based interaction we cannot exclude the possibility of the occurrence of technological breakdowns or their impact on the users. Furthermore, although the application was optimized to run above 80 frames per second (FPS) a possibility of episodic drops in frame rate is not excluded. To avoid any mismatch with participant's expectations of the appearance of their virtual bodies we have decided not to use an avatar representation. Instead, the only visual reference to their body was the virtual hands visualized through the Leap Motion controller. The gradual albeit unnatural appearance of the Disclaimers between the scenes could have affected the immersion. However, the participants were given sufficient time to adjust themselves to the interaction modality and the visualized hands.

Although Leap Motion itself is not immune from loss of tracking we have found that the healthy participants generally enjoyed the visualization of their hands which could

Table 5 Physiological variables during different scenes

Park1				Social			
Variable	Median	IQR	Group	Variable	Median	IQR	Group
CDA nSCR	4.33	5.25	a,b	CDA nSCR	8.50	5.42	b,c
CDA Tonic	1.12	0.87	a	CDA Tonic	2.06	0.86	a,b,c
HR	76.40	6.59	a	HR	77.69	8.16	a
RMSSD	35.34	9.29	b	RMSSD	31.75	10.72	b
Stroop				Social Stroop			
Variable	Median	IQR	Group	Variable	Median	IQR	Group
CDA nSCR	5.67	5.58	c	CDA nSCR	14.83	13.33	c
CDA Tonic	0.68	0.53	a,b,c	CDA Tonic	1.80	1.53	b
HR	84.88	13.76	b	HR	89.07	17.14	b
RMSSD	21.54	19.80	a	RMSSD	21.95	9.67	a
Park2							
Variable	Median	IQR	Group				
CDA nSCR	2.33	4.08	a				
CDA Tonic	1.39	1.36	c				
HR	73.79	9.72	a				
RMSSD	34.87	16.00	b				

Different letters in the “Group” column mean that a measured difference is statistically significant. Groups were assigned according to post hoc multiple comparisons of means (Tukey contrasts) at $p < 0.0125$ (Bonferroni corrected)

have had a beneficial effect on the sense of agency of the user. Healthy people have been used in this particular study which many have several implications for the interpretation and of the generalizability of the results. Namely healthy participants are less prone to effects that degrade interaction (frame-rate drops, loss of tracking) while on the other hand there can be hypothesized a larger effect on the patient population. This is one of the reasons such a limited but controlled intervention was used. There were reported instances of loss of hand tracking which would cause the virtual hands to appear in unnatural positions. This can be a serious matter of concern for patients. To address these issues further testing will include a measure of VR adverse effects.

7 Conclusions

While social stimuli can be manipulated to alter the level of arousal it may be inappropriate for the patient population, and would require prior validation. In this paper we have analyzed a safer way to manipulate the stress reaction, through the addition of the stroop task. Our findings indicate the utility of VR in inducing a stressful reactions through a combination of stressor for healthy subjects. Also, VR can serve as the bases for clinical VR research into a wide range of anxiety and somatoform disorders. Our first hypothesis was confirmed since both the VR Stroop test as

well as simulation of the social environment and their combination can elicit measurable stress reaction. While EDA measurements react better the stress caused by a combination of stressful task and social environment simulation, HR and RMSSD react more to other types of stress. Our second hypothesis was only confirmed in the case of CDA nSCR, HR and RMSSD. It appears that these three measurements change in time more than CDA Tonic. It is possible that the value of CDA Tonic influences values in subsequent scenes.

Implications for further work include the following. All four physiological measures should be used in further experiments because different measures likely reflect a different type of stress. Furthermore, there is no need to use a stressful task (i.e. Stroop test) alone as a stress condition. The optimal control stress condition to assess the capacity of social environment simulation to cause stress reaction is the combination of stressful task and social environment. Modification of the ordering of the scenes is not necessary. Moreover, the activity of virtual characters can be augmented with the addition of scripted conversations, with the aim of increasing the effectiveness of the exposure further. Effects on task performance of our implementation of VR Stroop test have not been assessed, and will be included in further experimentation.

Additional convenience in using the Stroop task is that the stress induction can be nuanced by altering the ratio of congruent and incongruent stimuli, or by manipulating the

time given for each answer. Great strides in graphics quality and continued search for interaction modalities convenient for VR continue to be made which raises the question on how will the stress induction procedures reflect the accelerating potential of VR. In this work, freely available resources were used to strengthen the prototyping capabilities in limited time dedicated to developing the VR stress induction. Looking ahead, researchers can expect a richer, more natural interaction, more believable renderings of environments and characters, and increasingly more available multimodal, safe and ergonomically appropriate systems.

Acknowledgements This work has been supported by the grant of Ministry of Health of Czech Republic 16-31457A. This study has been approved by the Ethical Committee at the Faculty of Medicine, Masaryk University, Brno, Czech Republic.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allen AP, Kennedy PJ, Dockray S, Cryan JF, Dinan TG, Clarke G (2017) The trier social stress test: principles and practice. *Neurobiol Stress* 6:113–126. <https://doi.org/10.1016/j.ynstr.2016.11.001> **(SI: Stressors in animals)**
- Armstrong CM, Reger GM, Edwards J, Rizzo AA, Courtney CG, Parsons TD (2013) Validity of the virtual reality stroop task (vrst) in active duty military. *J Clin Exp Neuropsychol* 35(2):113–123. <https://doi.org/10.1080/13803395.2012.740002>
- Benbow AA, Anderson PL (2019) A meta-analytic examination of attrition in virtual reality exposure therapy for anxiety disorders. *J Anxiety Disord* 61:18–26. <https://doi.org/10.1016/j.janxdis.2018.06.006> **(Virtual reality applications for the anxiety disorders)**
- Benedek M, Kaernbach C (2010) A continuous measure of phasic electrodermal activity. *J Neurosci Methods* 190(1):80–91. <https://doi.org/10.1016/j.jneumeth.2010.04.028>
- Blom KJ, Beckhaus S (2014) The design space of dynamic interactive virtual environments. *Virtual Real* 18(2):101–116. <https://doi.org/10.1007/s10055-013-0232-y>
- Calogiuri G, Litleskare S, Fagerheim KA, Rydgren TL, Brambilla E, Thurston M (2018) Experiencing nature through immersive virtual environments: environmental perceptions, physical engagement, and affective responses during a simulated nature walk. *Front Psychol* 8:2321. <https://doi.org/10.3389/fpsyg.2017.02321>
- Carl E, Stein AT, Levihn-Coon A, Pogue JR, Rothbaum B, Emmelkamp P, Powers MB (2019) Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. *J Anxiety Disord* 61:27–36. <https://doi.org/10.1016/j.janxdis.2018.08.003> **(Virtual reality applications for the anxiety disorders)**
- Cipresso P, Giglioli IAC, Raya MA, Riva G (2018) The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. *Front Psychol* 9:2086. <https://doi.org/10.3389/fpsyg.2018.02086>
- Critchley HD (2002) Electrodermal responses: what happens in the brain. *Neuroscientist* 8(2):132–142
- Diemer J, Alpers GW, Peperkorn HM, Shibani Y, Mühlberger A (2015) The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Front Psychol* 6:26. <https://doi.org/10.3389/fpsyg.2015.00026>
- Egan D, Brennan S, Barrett J, Qiao Y, Timmerer C, Murray N (2016) An evaluation of heart rate and electrodermal activity as an objective qoe evaluation method for immersive virtual reality environments. In 2016 eighth international conference on quality of multimedia experience (qomex) (p. 1–6)
- Frisch JU, Häusser JA, Mojzisch A (2015) The trier social stress test as a paradigm to study how people respond to threat in social interactions. *Front Psychol* 6:14. <https://doi.org/10.3389/fpsyg.2015.00014>
- Gutiérrez-Maldonado J, Alsina-Jurnet I, Rus-Calafell M (2009) Influence of personality and individual abilities on the sense of presence experienced in virtual environments. In Proceedings of the 2nd conference on human system interactions (pp. 149–152). IEEE Press
- Kellmeyer P (2018) Neurophilosophical and ethical aspects of virtual reality therapy in neurology and psychiatry. *Camb Q Healthc Ethics* 27(4):610–627
- Kelly O, Matheson K, Martinez A, Merali Z, Anisman H (2007) Psychosocial stress evoked by a virtual audience: relation to neuroendocrine activity. *CyberPsychol Behav* 10(5):655–662. <https://doi.org/10.1089/cpb.2007.9973>
- Kim K, Rosenthal MZ, Zielinski DJ, Brady R (2014) Effects of virtual environment platforms on emotional responses. *Comput Methods Programs Biomed* 113(3):882–893. <https://doi.org/10.1016/j.cmpb.2013.12.024>
- Kirschbaum C, Pirke K, Hellhammer D (1993) The “trier social stress test”: a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* 28(1–2):76–81
- Krejsa J, Kerous B, Liarokapis F (2018) A model for eye and head motion for virtual agents. In 10th international conference on virtual worlds and games for serious applications, vs-games 2018, wurzburg, Germany, September 5–7, 2018 (pp. 1–4). <https://doi.org/10.1109/VS-Games.2018.8493406>
- Lindner P, Miloff A, Zetterlund E, Reuterskiöld L, Andersson G, Carlbring P (2019) Attitudes toward and familiarity with virtual reality therapy among practicing cognitive behavior therapists: a cross-sectional survey study in the era of consumer vr platforms. *Front Psychol* 10:176. <https://doi.org/10.3389/fpsyg.2019.00176>
- Luks R, Liarokapis F (2019) Investigating motion sickness techniques for immersive virtual environments. In Proceedings of the 12th acm international conference on pervasive technologies related to assistive environments (pp. 280–288). New York, NY, USA: ACM. <https://doi.org/10.1145/3316782.3321535>
- Lupien SJ, Juster R-P, Raymond C, Marin M-F (2018) The effects of chronic stress on the human brain: from neurotoxicity, to vulnerability, to opportunity. *Front Neuroendocrinol* 49:91–105. <https://doi.org/10.1016/j.yfrne.2018.02.001>
- Maples-Keller J, Bunnell B, Kim S, Rothbaum B (2017) The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harvard Rev Psychiatry* 25(3):103–113. <https://doi.org/10.1097/HRP.0000000000000138>
- Norrholm SD, Jovanovic T, Gerardi M, Breazeale KG, Price M, Davis M (2016) Baseline psychophysiological and cortisol reactivity as

- a predictor of ptsd treatment outcome in virtual reality exposure therapy. *Behav Res Therapy* 82:28–37
- Pallavicini F, Cipresso P, Raspelli S, Grassi A, Serino S, Vigna C, Riva G (2013) Is virtual reality always an effective stressors for exposure treatments? Some insights from a controlled trial. *BMC Psychiatry* 13:3549–3557. <https://doi.org/10.1186/1471-244X-13-52>
- Parsons TD, Reinebold JL (2012) Adaptive virtual environments for neuropsychological assessment in serious games. *IEEE Trans Consum Electron* 58(2):197–204. <https://doi.org/10.1109/TCE.2012.6227413>
- Price M, Mehta N, Tone EB, Anderson PL (2011) Does engagement with exposure yield better outcomes? Components of presence as a predictor of treatment response for virtual reality exposure therapy for social phobia. *J Anxiety Disord* 25(6):763–770. <https://doi.org/10.1016/j.janxdis.2011.03.004>
- Rizzo A, Wiederhold M, Buckwalter J (1998) Basic issues in the use of virtual environments for mental health applications. *Stud Health Technol Inform*. <https://doi.org/10.3233/978-1-60750-902-8-21>
- Rizzo AA, Bowerly T, Buckwalter JG, Klimchuk D, Mitura R, Parsons TD (2009) A virtual reality scenario for all seasons: the virtual classroom. *CNS Spectr* 11(1):35–44. <https://doi.org/10.1017/S1092852900024196>
- Ruiz A, Peralta-Ramirez M, Garcia-Rios M, Muñoz M, Navarrete-Navarrete N, Blazquez-Ortiz A (2010) Adaptation of the trier social stress test to virtual reality: psycho-physiological and neuroendocrine modulation. *CyberTherapy Rehab* 3(4):405–415
- Schwartzman D, Segal R, Drapeau M (2012) Perceptions of virtual reality among therapists who do not apply this technology in clinical practice. *Psychol Serv* 9(3):310–315. <https://doi.org/10.1037/a0026801>
- Skarbez R, Brooks Jr FP, Whitton MC (2018a) Immersion and coherence in a stressful virtual environment. In *Proceedings of the 24th acm symposium on virtual reality software and technology* (p. 24)
- Skarbez R, Brooks FP Jr, Whitton MC (2018b) A survey of presence and related concepts. *ACM Comput Surv* 50(6):96
- Slater M (2009) Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos Trans R Soc Lond* 364(1535):3549–3557. <https://doi.org/10.1098/rstb.2009.0138>
- Slater M, Khanna P, Mortensen J, Yu I (2009) Visual realism enhances realistic response in an immersive virtual environment. *IEEE Comput Graphics Appl* 29(3):76–84. <https://doi.org/10.1109/MCG.2009.55>
- Slater M, Pertaub D, Barker C, Clark D (2006) An experimental study on fear of public speaking using a virtual environment. *CyberPsychol Behav* 9(5):627–633. <https://doi.org/10.1089/cpb.2006.9.627>
- Starcke K, Brand M (2012) Decision making under stress: a selective review. *Neurosci Biobehav Rev* 36(4):1228–1248. <https://doi.org/10.1016/j.neubiorev.2012.02.003>
- Stroop J (1935) Studies of interference in serial verbal reactions. *J Exp Psychol* 18(6):643–662
- Sutcliffe AG, Poullis C, Gregoriades A, Katsouri I, Tzanavari A, Herakleous K (2019) Reflecting on the design process for virtual reality applications. *Int J Hum Comput Interact* 35(2):168–179. <https://doi.org/10.1080/10447318.2018.1443898>
- Tarvainen MP, Niskanen J-P, Lipponen JA, Ranta-aho PO, Karjalainen PA (2014) Kubios hrv—heart rate variability analysis software. *Comput Methods Programs Biomed* 113(1):210–220. <https://doi.org/10.1016/j.cmpb.2013.07.024>
- Thayer JF, Lane RD (2000) A model of neurovisceral integration in emotion regulation and dysregulation. *J Affect Disord* 61(3):201–216
- Thayer JF, Lane RD (2009) Claude bernard and the heart-brain connection: further elaboration of a model of neurovisceral integration. *Neurosci Biobehav Rev* 33(2):81–88
- VanElzakker M, Dahlgren M, Davis F, Dubois S, Shin L (2014) From pavlov to ptsd: the extinction of conditioned fear in rodents, humans, and anxiety disorders. *Neurobiol Learn Mem* 113:3–18. <https://doi.org/10.1016/j.nlm.2013.11.014>
- Volonte M, Babu SV, Chaturvedi H, Newsome N, Ebrahimi E, Roy T, Fasolino T (2016) Effects of virtual human appearance fidelity on emotion contagion in affective inter-personal simulations. *IEEE Trans Visual Comput Graphics* 22(4):1326–1335. <https://doi.org/10.1109/TVCG.2016.2518158>
- Wu D, Courtney CG, Lance BJ, Narayanan SS, Dawson ME, Oie KS, Parsons TD (2010) Optimal arousal identification and classification for affective computing using physiological signals: virtual reality stroop task. *IEEE Trans Affect Comput* 1(2):109–118. <https://doi.org/10.1109/T-AFFC.2010.12>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.