This is the accepted manuscript of the article, which has been published in **Computers in human behavior**, 2020, 112, 106454.
https://doi.org/10.1016/j.chb.2020.106454

Eye contact in virtual reality – A psychophysiological study

Aleksi H. Syrjämäki[1], Poika Isokoski[2], Veikko Surakka[2], Tytti P. Pasanen[3], & Jari K. Hietanen[1]

[1] Human Information Processing Laboratory, Faculty of Social Sciences / Psychology, Tampere University

[2] Research Group for Emotions, Sociality, and Computing, Faculty of Technology and Communication Sciences, Tampere University

[3] Faculty of Social Sciences / Psychology, Tampere University

Address correspondence to Aleksi H. Syrjämäki, Human Information Processing Laboratory, Faculty of Social Sciences / Psychology, FI-33014, Tampere University, Tampere, Finland. E-mail: aleksi.syrjamaki@tuni.fi.

Declaration of interest. The authors report no potential conflicts of interest.

Acknowledgments. We would like to thank Samuli Pyssysalo and Katariina Sarlund for their help in data collection. We also thank Terhi Helminen, Mikko Peltola and Jonne Hietanen for their help in analyzing the data. This study was funded by Tampere University, a personal grant from the Emil Aaltonen Foundation to A.S., and the Emotions in Online Digital Media project funded by the Academy of Finland (grant #320766 to V.S.).

Abstract

This experiment investigated whether eye contact would evoke similar attention and emotion related psychophysiological responses in virtual reality (VR) as in a face-to-face interaction. Participants viewed a confederate in a live interaction (Live condition) and a confederate's avatar in VR (VR condition). In both conditions, the confederate / avatar was portraying direct and laterally averted gaze. Heart rate deceleration responses reflecting attention orienting were greater to direct gaze compared to averted gaze, and the effect was not significantly different between Live and VR conditions. However, skin conductance responses reflecting physiological arousal were larger in response to direct than averted gaze only in the Live condition. These results suggest that while eye contact with a live person evokes substantial attention and emotion related psychophysiological responses, the physiological effects of eye contact are diminished in VR.

Keywords: Eye gaze, arousal, attention, avatar, mediated communication

1. Introduction

Virtual reality (VR) technologies enable building applications that evoke experiences of being immersed in virtual environments. What separates VR from other digital media, such as traditional movies or video games, is the sense of presence, that is, a feeling of being in the virtual environment rather than observing it on a monitor (Witmer & Singer, 1998). This way VR resembles and can simulate real life circumstances better than stimulations in 2D. As VR is developing fast for various purposes (e.g. telecommunication and entertainment), it will likely develop further towards a medium for human-human interaction. Understanding of human social-emotional behavior and processes in VR is important as such, but also for developing more natural and rich interactions for this new medium. Already now, virtual social interactions can be enhanced with motion trackers and eye trackers allowing users to communicate with each other using realistic gestures, such as hand and eye movements (cf. Dodds, Mohler, & Bühlthoff, 2011; Pejsa, Andrist, Gleicher, & Mutlu, 2015).

Facially communicated information is centrally important in real life social interactions. Gaze direction is especially interesting for VR applications, as VR headsets begin to have gaze direction tracking as an additional feature. In social interactions, we infer important information such as the interaction partner's intentions, motivations, and direction of attention based on their gaze direction (for a review, see Kleinke, 1986). When another person is averting their eye gaze, it can indicate, for instance, that there is something important in the gazed-at location (Frischen, Bayliss, & Tipper, 2007), or that the person wants to avoid social interaction (Adams & Kleck, 2005). Another person's direct gaze (gaze directed at the perceiver's eye region), instead, signals that the other person's attention is directed to the perceiver.

Due to its high social significance, direct gaze evokes various attention and emotion related responses in the perceiver (for reviews, see Conty, George, & Hietanen, 2016; Hietanen, 2018; Senju & Johnson, 2009). From an early age, people preferentially orient their attention towards faces that are portraying direct gaze (Farroni, Csibra, Simion, & Johnson, 2002; Senju & Johnson, 2009). There is evidence that direct gaze “captures” the perceiver’s attention rapidly and automatically (e.g., Conty, Tijus, Hugueville, Coelho, & George, 2006; Lyyra, Astikainen, & Hietanen, 2018; von Grünau & Anston, 1995). At the physiological level, attention orienting towards affectively and socially salient stimuli is typically associated with a brief deceleration of heart rate (HR; Graham & Clifton, 1966; Kreibig, 2010; Öhman, Hamm, & Hugdahl, 2000), and indeed, it has been reported that this response is more pronounced after perceiving direct gaze than averted gaze (Akechi et al., 2013; Myllyneva & Hietanen, 2015b; Wieser, Pauli, Grosseibl, Molzov, & Mühlberger, 2010).

Another person’s direct gaze has also been shown to evoke emotional reactions. In a recent review, Hietanen (2018) noted that studies measuring subjective judgments of emotions have reported mixed findings, with some reporting direct gaze (vs. other gaze directions) to evoke more positive affect, others reporting an opposite effect, and others finding no effects at all. However, studies measuring implicit affective reactions – reactions that are relatively automatic and free from higher level cognitive evaluation – have provided a clearer picture, suggesting that another person’s direct gaze increases arousal and positive affect. One particularly consistent finding is that, at the physiological level, direct gaze relative to averted gaze evokes larger skin conductance responses (SCR; e.g., Helminen, Kaasinen, & Hietanen, 2011; Hietanen, Leppänen, Peltola, Linna-aho, & Ruuhiala, 2008; Jarick & Bencic, 2019; Nichols & Champness, 1971). SCRs reflect activity of the autonomic sympathetic nervous system (Critchley, 2002; Dawson, Schell, & Filion, 2000). Autonomic activation is one of the components of the emotional compound response (Plutchik, 1980), thus these studies demonstrate that direct gaze increases affective arousal (for converging evidence using other physiological measurements, see Gale, Lucas, Nissim, & Harpham, 1972; Gale, Spratt, Chapman, & Smallbone, 1975).

Interestingly, these attention and emotion related physiological responses to direct gaze seem to be triggered by live faces, but not by viewing face photographs or videos. Several early eye contact studies reported more physiological arousal during an interaction with a live confederate portraying direct gaze compared to other gaze directions (as measured by SCRs, HR acceleration and EEG alpha activity; Gale et al., 1972, 1975; Kleinke & Pohlen, 1971; Nichols & Champness, 1971). In contrast, numerous subsequent studies found no evidence that direct gaze on pictures or videos would evoke larger physiological responses than other gaze directions (measured by SCRs, HR responses and pupil size; Donovan & Leavitt, 1980; Joseph, Ehrman, McNally, & Keehn, 2008; Kampe, Frith, & Frith, 2003; Leavitt & Donovan, 1979; Lyyra, Myllyneva, & Hietanen, 2018; Schrammel, Pannasch, Graupner, Mojzisch, & Velichkovsky, 2009; Wieser, Pauli, Alpers, & Mühlberger, 2009). These differing findings have been suggested to be explained by differences in the stimuli used in the studies (live confederates vs. face pictures; e.g., Hietanen et al., 2008). Evidence supporting this interpretation was provided in experiments directly comparing the effects of live and pictorial direct gaze (Hietanen et al., 2008; Pönkänen, Alhoniemi, Leppänen, & Hietanen, 2011; Pönkänen, Peltola, & Hietanen, 2011; Prinsen & Alaerts, 2019). In these studies, participants viewed a live confederate's face through an electronic shutter window, and in another presentation condition, photographs or videos of the same face. It was found that only the live face's direct gaze, relative to other gaze directions, evoked larger physiological and neural responses (SCRs, frontal EEG asymmetry and event related potentials), but the pictorial gaze direction had no effects on the responses. These results indicate that a reciprocal eye contact with a live person evokes different psychophysiological responses than a mere perception of a facial image with eyes directed towards oneself.

The difference in responses to live versus pictorial faces likely reflects the perceiver's sense of being seen by another person in the live situation. In a study by Myllyneva and Hietanen (2015b), participants viewed live faces portraying direct and averted gaze, and their experience of being seen was manipulated. In one presentation condition, a bogus half-silvered mirror was placed between the participant and the live model, so that it supposedly blocked the model person from seeing the participant while still allowing the participant to see the model. In another presentation condition, there was no half-silvered mirror between them, and the participant was aware the model could also see them. Importantly, the participant's view of the model was identical in these two conditions. HR deceleration responses, P3 event-related brain potentials (reflecting attention orienting), and SCRs (reflecting physiological arousal) were measured. The responses were greater to direct gaze as compared to averted gaze, but only in the condition where participants believed that the model could see them. In the condition where participants believed the model's view was blocked, the model's gaze direction had no effects on the responses. The results indicate that the experience of being the target of another's attention is a necessary component in direct gaze eliciting these psychophysiological responses. However, the experience of being seen does not seem to be solely driving these effects. In a related study, the knowledge of being looked at did not evoke the enhanced HR deceleration responses and larger SCRs in a condition where participants themselves could not simultaneously see the looker (manipulated by leading the participants to believe that a half-silvered mirror blocked their but not the model's view; Myllyneva & Hietanen, 2015a). These findings suggest that in order for direct gaze to magnify physiological attention orienting and arousal responses, the perceiver must see the person portraying direct gaze and simultaneously have a sense of being seen by the person.

Importantly, eye contact in VR with another user's avatar (the user's virtual representation) can create an impression of this kind of bidirectional visual interaction between two people. Previous research has shown that even in technologically mediated social interactions, people engage in mentalization processes, inferring what their interaction partners experience (Gallagher, Jack, Roepstorff, & Frith, 2002; Kircher et al., 2009). Thus, eye contact with an avatar can likely evoke an experience of being seen, similarly to eye contact with a live person. However, no previous study has investigated whether the psychophysiological responses to virtual eye contact would be similar, or possibly diminished, when compared to real eye contact. Previous research has found that direct gaze evokes larger HR deceleration responses than averted gaze, both in live interaction (Akechi et al., 2013; Myllyneva & Hietanen, 2015b), and in VR (Wieser et al., 2010). In the Wieser et al. study, participants wore VR headsets and were placed in a virtual environment, in which virtual characters walked towards them. It was found that HR decelerated more when the character was portraying direct gaze, as compared to averted gaze. From these findings, it could be postulated that eye contact in VR indeed evokes similar psychophysiological responses as live eye contact, although differences between studies of course make it impossible to draw far-reaching conclusions. There are also important distinctions between live and virtual interactions, which could be expected to attenuate effects of direct gaze in VR. In virtual eye contact, the individuals do not see each other's real faces, but rather their avatars. Due to this, social presence (the sense of being in the presence of another person; Heeter, 1992) might be lower, possibly making virtual eye contact a less salient experience than natural eye contact. Consistent with this notion, prior research has reported lower physiological arousal during speaking to a virtual vs. live audience (Owens & Beidel, 2015), and lower subjective arousal in response to virtual humans vs. live people invading the personal space (Wilcox, Allison, Eflassy, & Grelik, 2006).

The aim of the present study was to test whether eye contact in VR would evoke similar attention and emotion related psychophysiological responses as eye contact in a live interaction. Participants were shown faces in two different blocks. In one block, they viewed a live confederate's face (Live condition). In another block, participants wore a VR head-mounted display (HMD) and viewed an avatar in a virtual environment (VR condition). Critically, participants were led to believe that the avatar was controlled by the confederate who was also wearing a VR HMD. Participants were led to believe that both HMDs were equipped with eye trackers, allowing the participant and the confederate to see each other's gaze directions. In both blocks, we measured participants' HR and skin conductance responses to the confederate's / avatar's face portraying direct and laterally averted gaze. We expected that direct gaze would evoke larger attention orienting responses (as measured by HR deceleration) and physiological arousal responses (SCRs) than averted gaze, in both Live and VR conditions (e.g., Akechi et al., 2013; Wieser et al., 2010). Because previous research has reported lower subjective and physiological responses to virtual than live people (Owens & Beidel, 2015; Wilcox et al., 2006), we also hypothesized that the effects could be smaller in magnitude in the VR condition than in the Live condition. The physiological measurements were complemented by questionnaires, measuring participants' subjective feelings in response to different gaze directions in the Live and VR conditions as well as their feelings of social presence during the Live and VR blocks.

2. Materials and methods

2.1. Participants

We gathered data from 40 participants. This exceeds the required sample size for finding a medium effect (d = 0.50) at 0.80 power and α level of 0.05 (Cohen, 1992; Faul, Erdfelder, Lang, & Buchner, 2007). Because the effect of live direct vs. averted gaze on SCRs is large (reported *d*s ranging from 0.70 to 0.98; Hietanen et al., 2018; Myllyneva & Hietanen, 2015b), we thus expected statistical power to be sufficient for detecting the effect even if it was diminished in the VR condition (i.e., a medium effect of d = 0.50). Participants were required to be at least 15 years old and to not have diagnosed psychiatric or neurological disorders. One participant was excluded from all analyses due to expressing awareness of the cover story of the experiment (see below for details), and another due to reporting suffering from a psychiatric disorder. Thus, the analyzed sample consisted of 38 participants, aged 16–53 (M = 23.6, SD = 9.4, 12 males, 26 females). Two additional participants (a male and a female) were excluded from the analyses of the HR data because of technical problems in the HR recording. All participants signed an informed consent form, and they were rewarded either with partial course credit or with a movie ticket. An ethical statement for the study was obtained from the Ethics Committee of the Tampere region.

2.2. Apparatus and stimuli

For an illustration of the experimental setting, as well as the stimulus faces in the VR condition, see Figure 1. In the Live condition, the stimulus faces were presented through a 21.5 cm × 38 cm voltage-sensitive liquid crystal window (LC window; NSG UMU Products Co., Ltd., Ichihara, Chiba, Japan), which could be changed between opaque and transparent states. The LC window was operated by E-Prime 2.0 software (Psychology Software Tools, Inc., Pittsburgh, Pennsylvania, USA). The LC window was attached to a black partition which was placed on a table between the participant and the confederate. For the VR condition, we created a 5 m × 5 m virtual experiment room, which was equipped with a (virtual) LC window resembling the real LC window. On one side of the virtual LC window, there was an avatar (see below for details), but otherwise the virtual room was empty. In the VR block, the stimuli and the environment were presented with a Samsung Odyssey VR HMD (Samsung Electronics Co., Ltd., Suwon, South Korea), which tracked the participant’s head orientation and position, and displayed stereoscopic image with a resolution of 1440 × 1600 per eye at a refresh rate of 90 Hz. The program for the VR condition was created with Unreal Engine 4.19.2. (Epic Games, Inc., Cary, North Carolina, USA), and was run on a desktop computer.

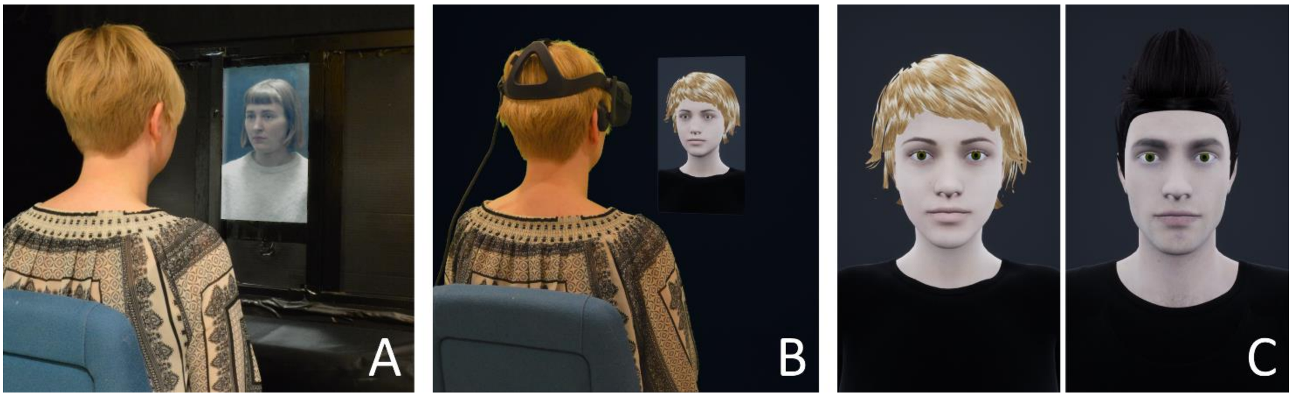

Figure 1. A. An illustration of the experimental setting in the Live condition. The participant and the confederate were seated on opposite sides of an LC window. When the LC window was turned transparent for three seconds, the participant saw the confederate who was portraying either direct or averted gaze. **B.** An illustration of the experimental setting in the VR condition. The participant was wearing an HMD, and saw a virtual room, similar to the real laboratory where they were physically seated. Behind a virtual LC window, there was an avatar, ostensibly representing the confederate. In the figure, the participant is shown for illustrative purposes only, as they were not actually visible in the VR environment. **C.** The two avatars which were used as stimuli in the VR condition.

In the Live condition, a confederate acted as a stimulus person. For the VR condition, we created two stimulus avatars with the 3D animation software DAZ Studio (Daz Productions, Inc., Salt Lake City, Utah, USA). The confederate's / avatar's gender matched the participant's gender. In both conditions, the face stimuli were presented against a dark background at an approximate distance of 50 cm from the participant. The confederate maintained a neutral expression with a slight muscle tonus in the lower part of the face to avoid a sullen expression, avoided eye blinks, and stayed relatively motionless during the stimulus presentation. Similarly, the avatar in the VR condition had a neutral facial expression, made no eye blinks, and stayed still, except for slight vertical movement in the upper part of the body to create an impression of breathing. Stimulus presentation was controlled by alternating the opacity of the LC window (real and virtual) between opaque and transparent states. On each trial, the stimulus face was shown for three seconds at a time. The confederate's / avatar's eyes were directed either at the participant's eyes (direct gaze) or 20° to the left or right (averted gaze). The experimenter controlled the avatar's gaze direction using a keyboard.

2.3. Procedure

The experiment was run by a male experimenter and two research assistants (a male and a female). The research assistant whose gender matched the participant's gender acted as a confederate, pretending to be another participant. The participant arrived to the laboratory with the confederate. They were told that the experiment was investigating physiological responses in social interactions, and that they would be interacting with each other in two tasks. The order of the blocks (Live condition / VR condition) was counterbalanced across participants[1]. At the beginning, instructions were only given for the first block (see below for details) so that knowledge about the second block would not influence participants' responding during the first block. To ensure that participants would not figure out the second block, the equipment used in the latter block (LC window / HMDs) were hidden from view. Detailed instructions about the task were given to the participants in a separate room to lead them to believe that instructions were simultaneously also given to the confederate. In these instructions, participants were told that they and the confederate would be seated on opposite sides of an LC window, and that they would see each other when the LC window was transparent. Participants were instructed to sit still and look at the confederate's face when the LC window was transparent. Upon returning to the laboratory, the electrodes for the physiological measurements were attached (see below for details). To maintain the cover story, electrodes were also attached to the confederate, but no measurements were actually taken. In both conditions, the participant and the confederate were seated on opposite sides of the partition during the task.

In preparation for the Live block, participants were introduced to the functioning of the LC window. To make the Live condition include a similar break during the preparation phase as the VR condition (see below), participants were asked to wait for approximately one minute after being introduced to the task and the LC window, supposedly to let their HR stabilize. The LC window was then made transparent, and the participant's and the confederate's head position were adjusted to make sure they were vertically and horizontally aligned. The experimental block was then started.

In preparation for the VR block, the participant and the confederate were introduced to the HMDs. They were told the HMDs were equipped with eye trackers, which allowed the participant and the confederate to see each other's gaze directions. The participant and the confederate were given a description of the virtual room they would be in, and to familiarize the participant with the stimulus character, they were shown a picture of both the participant's and the confederate's avatar. During the VR block, the participant and the confederate were physically seated at the same positions as in the Live block, i.e., on opposite sides of the partition. Before beginning the task, the participant put on the HMD, which was then adjusted to make the image as sharp as possible. The confederate also put on an HMD, but in reality, this device was not connected to a computer, and the confederate was only pretending to be in the same virtual environment with the participant. After putting on the HMD, the participant was first seated alone in a 5 m × 5 m virtual room containing some furniture to look at. The room was different from the virtual room where the task took place, and its purpose was to give the participant a space in which to get acclimated to VR. Participants were instructed to look around freely so that they would get used to VR, and supposedly to give time for their HR to stabilize. After approximately one minute, the experimenter told that the eye trackers would be calibrated. Participants were instructed to fixate on five dots, which appeared successively in the participant's field of view. The faux calibration process was ostensibly done first to the confederate, and then to the participant. After this, the participant, and ostensibly the confederate, was moved to the virtual experiment room. The participant was placed in the room so that the stimulus avatar, on the opposite side of the virtual LC window, was vertically and horizontally aligned with the participant's view. Next, the function of the eye trackers was supposedly tested by briefly turning the virtual LC window transparent a few times and asking the participant and the confederate to look in different directions and to check that they perceived each other's gaze directions correctly. The aim was to make the participant believe that they and the confederate could see each other's gaze directions, and additionally, to further familiarize the participant with the stimulus character. After this, the experimental block was started.

During both experimental blocks, participants were presented with the confederate's or the avatar's face 16 times (8 times with direct gaze, 8 times with averted gaze, equally to the left and the right). The trial order was pseudorandomized so that there were no more than two trials with direct or averted gaze in succession. The inter-stimulus interval was at least 10 s, or until the participant's skin conductance had stabilized after the previous trial.

After both blocks, we administered questionnaires. First, we measured participants' subjective feelings evoked by the gaze directions. Participants were again shown the face stimuli (with direct / averted gaze), both times followed by the self-assessment manikin, a questionnaire measuring the subjective experience of arousal (e.g., feeling calm vs. aroused) and valence (e.g., feeling unpleasant vs. pleasant) on a 1-9 scale (Bradley & Lang, 1994). Before showing the stimuli, the participant was introduced to the questionnaire, while the confederate was asked to read written instructions about what to do in this phase. The LC window was then made transparent twice for three seconds, revealing the stimulus face portraying direct or averted gaze (randomly to the left or right). The order of the direct and averted gaze presentations was counterbalanced across participants, and the order was the same in both Live and VR conditions. After each presentation, participants were asked to assess their experience of arousal and valence during the presentation. As a manipulation check after each presentation, participants also indicated on a 1-9 scale (strongly disagree – strongly agree) whether the other person / avatar was looking at him / her. In both conditions, the answers to these items were given orally because in the VR condition, participants could not use a pen while wearing the HMD and they had no input devices besides the HMD. In the VR condition, the HMDs were then taken off after responding to these questionnaires.

After completing both of these stimulus presentations and questionnaires, the participant and the confederate then filled in the Social presence form (SPF; Short, Williams, & Christie, 1976) measuring their sense of social presence during the experimental block. The standard version of the SPF was expanded with the following extra items: “I didn’t feel the other person was socially present at all – I strongly felt the other person was socially present”, “I didn’t feel the other person was physically present at all – I strongly felt the other person was physically present”, and “I was not absorbed in the situation at all – I was strongly absorbed in the situation”. Cronbach’s αs for the expanded SPF in the Live block and the VR block were .88 and .81, respectively, suggesting good internal consistency. Participants then assessed the distance to the other person in centimeters. After this, they indicated the amount of nausea they experienced during the block on a 1-9 scale (not at all – very much).

After completing the second block (Live or VR), the participant and the confederate filled in two more questionnaires. First, they indicated how often they had used VR devices on a 1-9 scale (never – daily) and listed the types of VR applications they had used. However, these data were not used in the analyses, because there was very little variation in the responses, with almost all participants reporting having no or very little experience with VR. We then probed whether participants had inferred the cover story in the VR condition. The participant and the confederate were presented with two consecutive questions, to which they were asked to write their answers: “What do you think the experiment was about?”, and “Do you think there was something the experimenter did not tell you about? If yes, what was it?”. As mentioned above, one participant was excluded from all analyses due to expressing awareness of the cover story. Finally, participants were thoroughly debriefed.

2.4. Acquisition of the physiological data

For the HR measurements, we recorded electrocardiogram (ECG) with two electrodes (Ag / AgCl) coated with electrode gel, placed on the participants' anterior forearms. For the skin conductance measurements, two electrodes (Ag / AgCl) coated with isotonic electrode paste were attached to the palmar surface of the distal phalanges of the index and middle fingers of the participants' non-dominant hand. A ground electrode was placed on the opisthenar of the non-dominant hand. The sampling rate for the digitized signals was 1000 Hz.

2.5. Analysis of the physiological data

Heart rate. The attention orienting response consists of rapid deceleration of HR, followed by acceleration towards the baseline (Graham & Clifton, 1966). To analyze this response, we used an analytic strategy used in several earlier studies (e.g., Akechi et al., 2013; Kolassa & Miltner, 2006; Richards & Turner, 2001). In this analysis, HR is calculated for multiple short time intervals following stimulus presentation, making it possible to distinguish momentary variations in HR.

From the ECG data, the intervals between two successive R-waves (inter-beat interval, IBI) were measured using an in-house Matlab-based algorithm (Peltola, Hietanen, Forssman, & Leppänen, 2013). R-peaks were first detected using the algorithm. Then incorrectly detected and missing R-peaks were corrected manually. For a period between 1 s pre-stimulus (baseline) and 6 s post-stimulus within each trial, the IBIs were quantified and assigned to 1-s intervals by weighting each IBI by the proportion of the 1-s interval occupied by that IBI (Richards

& Turner, 2001). Trials, on which one or more R-peaks within the 1-s baseline could not be reliably detected, were removed from further analyses (0.1 % of trials). The 1-second post-stimulus intervals, on which one or more R-peaks could not be reliably detected, were also removed from the analyses (0.1 % of remaining intervals). The IBIs were converted to beats per minute (BPM), and a change score was then calculated for each of the accepted post-stimulus time intervals by subtracting the baseline from the respective BPM value. The change scores were then averaged across trials for each condition.

Skin conductance responses. The skin conductance data were re-sampled offline to 100 Hz and filtered with a 10 Hz low-pass filter. The skin conductance response (SCR) was defined as the maximum amplitude increase from the preceding minimum within the analysis window of 0.9-6.5 s from the stimulus onset. This analysis window was chosen because an SCR is evoked 0.9-3.5 s after the presentation onset of a visual stimulus (Sjouwerman & Lonsdorf, 2019) and typically takes up to 3 s to reach its peak (Dawson et al., 2000). If there were two or more peaks within the analysis window, the maximum amplitude was calculated from the first peak. A trial was coded as a zero response if there was no peak (e.g., if there was a steady amplitude increase within the analysis window), if the maximum amplitude increase was less than 0.01 μS, or if the amplitude increase started later than 3.5 s after the stimulus onset. If there was an amplitude increase of 0.01 μS or more within 0.9 s from the stimulus onset, the trial was rejected (15.6 % of all trials). For each participant, the data from accepted trials, including trials with zero responses, were then averaged for each condition. This results in the magnitude of the skin conductance response, a measure that combines response size and response frequency (e.g., Dawson et al., 2000). To reduce the skewness of the resulting distribution, a cubic root transformation was conducted for the values used in the statistical analyses. For the sake of clarity, we present untransformed values in the results section below.

3. Results

3.1. Physiological data

Heart rate. Mean HR change scores in each analyzed time interval in each condition are presented in Figure 2A. The HR change scores were subjected to a 2 (Presentation mode: Live / VR) × 2 (Gaze direction: Direct / Averted) x 6 (Time interval: 0-1s, 1-2s,…, 5-6s) repeated-measures ANOVA. Greenhouse-Geisser correction was used when the assumption of sphericity was violated. There was a significant main effect of Time interval ($F(2.2, 78.2) = 12.63$, $p < .001$, $\eta_p^2 = .27$, Greenhouse-Geisser corrected); HR decelerated shortly after the stimulus presentation reflecting orienting of attention, and then accelerated towards the baseline. Two interactions were also significant. First, there was a Presentation mode × Time interval interaction ($F(3.0, 106.6) = 2.93$, $p = .036$, $\eta_p^2 = .08$, Greenhouse-Geisser corrected). We conducted follow-up t-tests to break down this interaction and found that HR was lower in the Live condition than in the VR condition in the last two time intervals (both *p*s < .05 without correction for multiple comparisons, see Figure 2B). This suggests that the attention orienting response was larger in response to live faces than avatar faces. Critically, the ANOVA also revealed a significant Gaze direction × Time interval interaction ($F(2.8, 96.5) = 3.09$, $p = .034$, $\eta_p^2 = .08$, Greenhouse-Geisser corrected). This interaction suggests that direct gaze evoked larger HR deceleration than averted gaze did: Follow-up t-tests showed that HR was lower in response to direct vs. averted gaze in the 3-4s time interval ($p = .038$ without correction for multiple comparisons, see Figure 2C). Importantly, there was no Presentation mode × Gaze direction × Time interval interaction ($F(2.3, 80.8) = 0.61$, $p = .571$, $\eta_p^2 = .02$, Greenhouse-Geisser corrected), or Presentation mode × Gaze direction interaction ($F(1, 35) = 2.19$, $p = .148$, $\eta_p^2 = .06$) showing that the effect of gaze direction on HR deceleration did not differ significantly between VR and Live conditions.

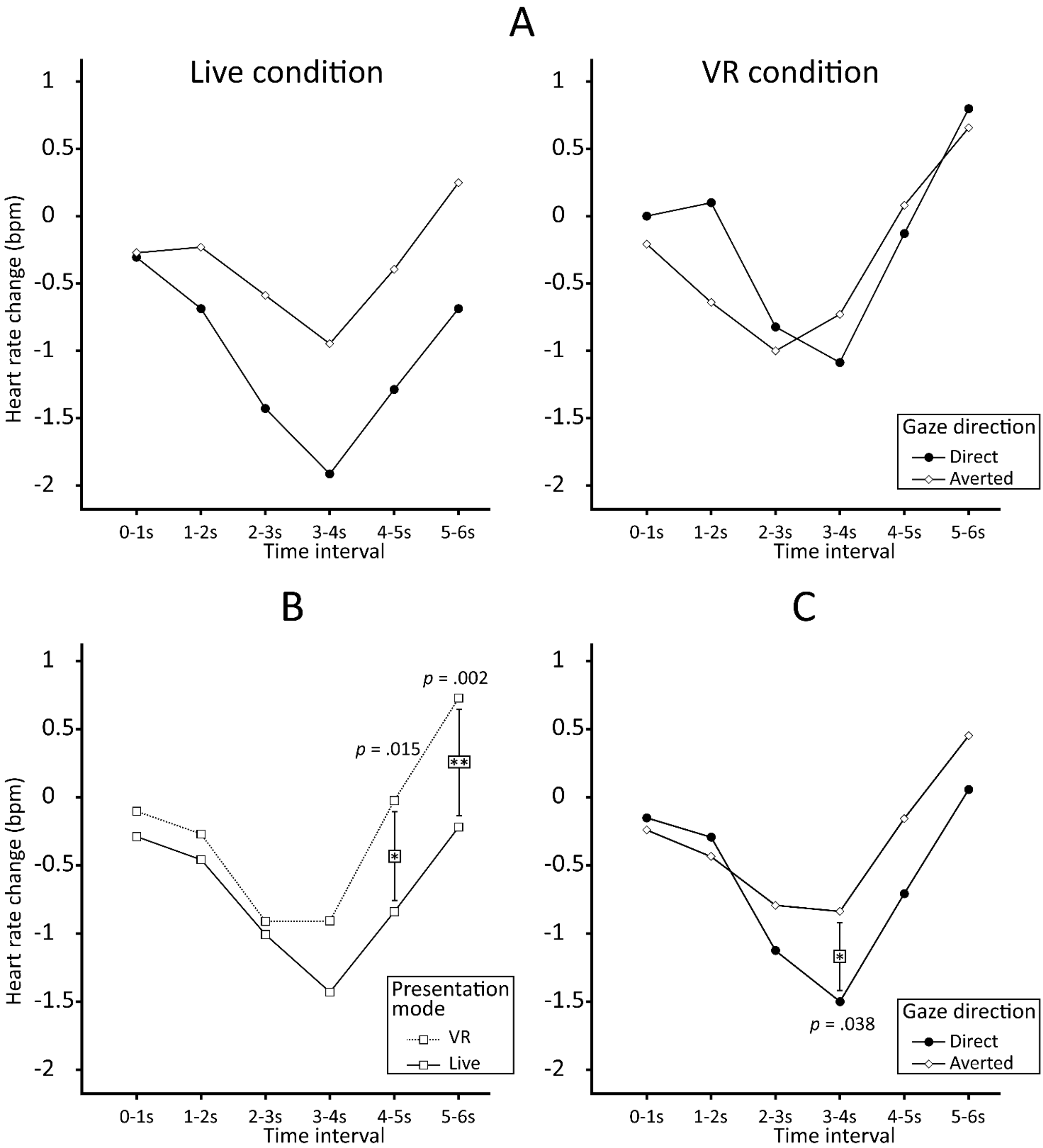

Figure 2. Mean HR change scores in different conditions at each time interval. In upper part of the Figure, mean HR change scores at each time interval are presented as a function of Presentation mode and Gaze direction (A). In lower part of the Figure, the results are presented in the Live and VR conditions averaged over the two gaze directions (B), and in the direct and averted gaze conditions averaged over the two presentation conditions (C). In B and C, the significant differences

between conditions, as indicated by paired-samples t-tests, are pointed out and the respective p-values are displayed. * $p < .05$, ** $p < .01$.

While the difference between the conditions was not statistically significant, a visual inspection of the data (Figure 2A) suggests that the effect of gaze direction in the VR condition may have been extremely small. It is not possible to determine whether an effect is zero, but the Two One-Sided Test (TOST) procedure allows testing whether an effect is statistically equivalent to zero (i.e., significantly smaller than what would be considered a meaningful effect; Lakens, 2017; Lakens, Scheel, & Isager, 2018). We used this procedure to compare HR deceleration in response to direct and averted gaze within both presentation conditions, averaged over the time intervals. Equivalence bounds were set at a medium effect size ($d = \pm 0.50$). As is expected based on the results of the ANOVA, paired-samples t-tests showed that in neither condition the differences between direct and averted gaze were statistically significant ($p = .106$ and $p = .700$ for the Live and VR conditions, respectively). The equivalence test for the Live condition was non-significant ($t(35) = 1.34$, $p = .094$). Most importantly, the equivalence test showed that in the VR condition, the effect of gaze direction was statistically equivalent to zero ($t(35) = 2.61$, $p = .007$). Therefore, we can conclude that in the VR condition, gaze direction did not have a meaningfully large effect on HR deceleration.

Skin conductance responses. Mean SCRs in each condition are presented in Figure 3. The SCR data were analyzed using a 2 (Presentation mode: Live / VR) × 2 (Gaze direction: Direct / Averted) repeated-measures ANOVA. The analysis revealed main effects of both Presentation mode ($F(1, 37) = 11.07$, $p = .002$, $\eta_p^2 = .23$) and Gaze direction ($F(1, 37) = 15.35$, $p < .001$, $\eta_p^2 = .29$). Most importantly, the main effects were qualified by a Presentation mode × Gaze direction interaction ($F(1, 37) = 10.16$, $p = .003$, $\eta_p^2 = .22$). Follow-up t-tests showed that direct gaze evoked larger SCRs than averted gaze in the Live condition ($t(37) = 5.70$, $p < .001$, $d = 0.95$), but not in the VR condition ($t(37) = 0.09$, $p = .932$, $d = 0.01$). The TOST procedure performed at equivalence bounds of $d = \pm 0.50$ demonstrated that the effect of direct vs. averted gaze on SCRs in the VR condition was equivalent to zero ($t(37) = 3.00$, $p = .002$). When comparing the presentation modes, SCRs evoked by direct gaze were larger in the Live condition than in the VR condition ($t(37) = 4.31$, $p < .001$, $d = 0.78$). The difference between Live and VR conditions in SCRs evoked by averted gaze was not statistically significant ($t(37) = 1.71$, $p = .096$, $d = 0.30$), but not statistically equivalent to zero ($t(37) = 1.37$, $p = .089$).[2]

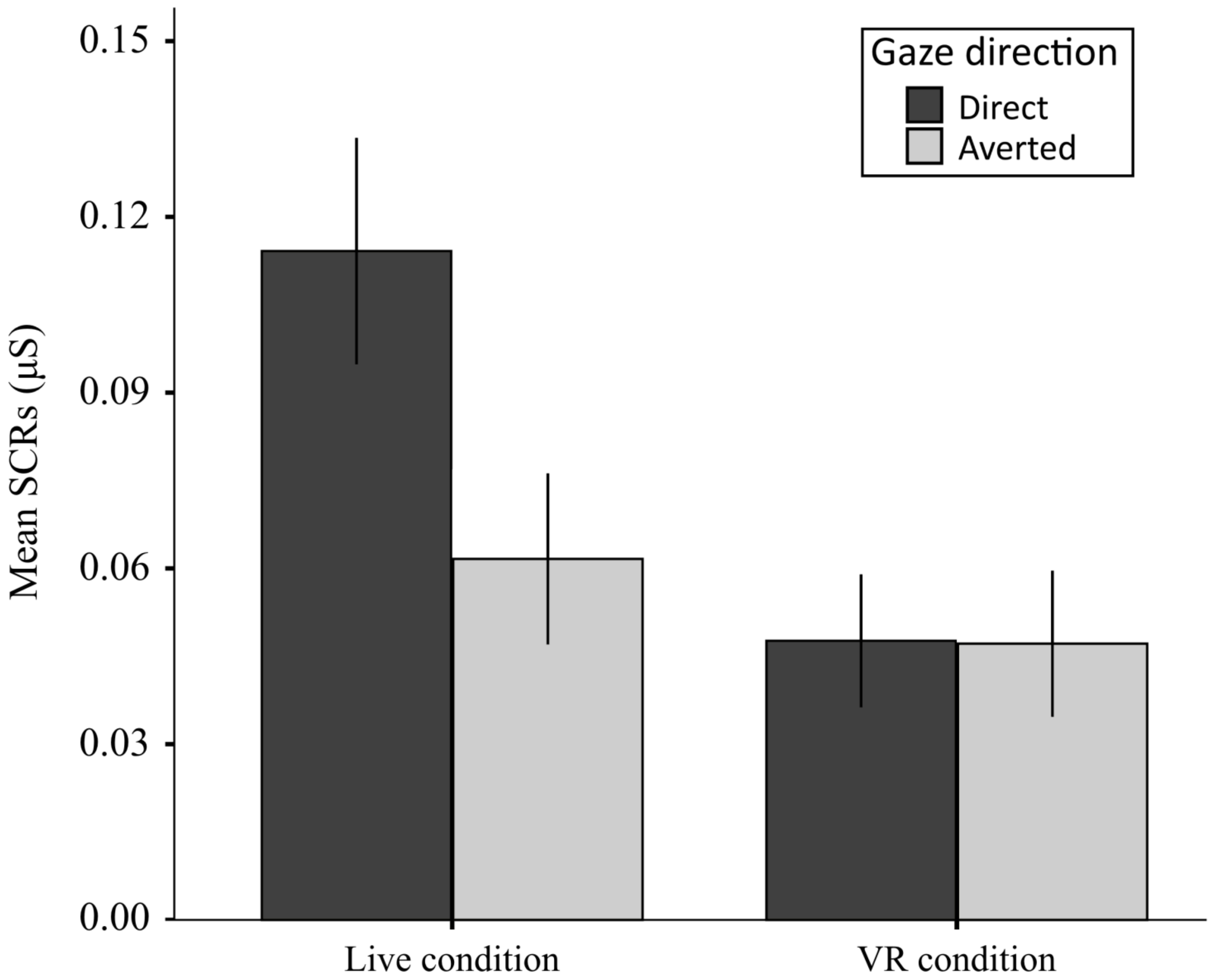

Figure 3. Mean SCRs in response to direct and averted gaze in the two presentation conditions. Error bars denote standard error of the means.

3.2. Questionnaire data

Responses to gaze directions. Means and standard deviations for the manipulation check item, and subjective arousal and valence ratings in each condition are presented in Table 1. These data were analyzed with 2 (Presentation mode: Live / VR) × 2 (Gaze direction: Direct / Averted) repeated-measures ANOVAs. For the manipulation check item, there was a significant main effect of Gaze direction ($F(1, 37) = 1275.4$, $p < .001$, $\eta_p^2 = 0.97$). The agreement that the face was looking towards oneself was larger for direct gaze stimuli than averted gaze stimuli ($M_{\text{direct}} = 8.38$, $SD = 1.10$; $M_{\text{averted}} = 1.17$, $SD = 0.72$). Importantly, the effect of Presentation mode or the interaction between Gaze direction and Presentation mode were not significant (ps $\geq .79$). For arousal ratings, there was a significant main effect of Gaze direction ($F(1, 37) = 9.87$, $p = .003$, $\eta_p^2 = 0.21$), showing that arousal was judged higher in response to direct than averted gaze ($M_{\text{direct}} = 2.80$, $SD = 1.48$; $M_{\text{averted}} = 2.22$, $SD = 1.21$). The effect of Presentation mode or the interaction between Gaze direction and Presentation mode were not significant for arousal ratings (ps $> .19$). For valence ratings, no main effects or interactions were found (ps $> .43$).

Table 1. Means and standard deviations for the manipulation check item (agreement on whether the face was looking towards oneself), and subjective arousal and valence ratings in each condition.

	Manipulation check		**Arousal**		**Valence**	
Condition	Direct	Averted	Direct	Averted	Direct	Averted
Live	8.39 (1.15)	1.21 (1.30)	2.97 (1.72)	2.18 (1.23)	5.58 (1.67)	5.45 (1.84)
VR	8.37 (1.68)	1.13 (0.67)	2.63 (1.57)	2.26 (1.43)	5.53 (1.62)	5.63 (1.87)

Responses to presentation modes. To compare participants' experiences of the two presentation conditions, the average SPF scores, nausea ratings and judgments of the distance to the stimulus person were analyzed with paired-samples t-tests. Nausea and distance ratings were missing from one male participant. Participants reported feeling more social presence in the Live condition than in the VR condition (M_{live} = 4.04, *SD* = 1.06; M_{VR} = 3.21, *SD* = 0.90; $t(37) = 4.85$, $p < .001$, $d = 0.73$), and more nausea in the VR condition than in the Live condition (M_{live} = 1.43, *SD* = 1.12; M_{VR} = 2.11, *SD* = 1.56; $t(36) = 2.83$, $p = .008$, $d = 0.58$), likely due to simulation sickness in VR (Sharples, Cobb, Moody, & Wilson, 2008). The distance to the face was estimated to be similar in the two conditions (M_{live} = 58.0 cm, *SD* = 21.6; M_{VR} = 56.1 cm, *SD* = 29.6; $t(36) = 0.48$, $p = .636$, $d = 0.10$), suggesting that the VR headset depicted distances accurately, unlike some older devices on which distances were underestimated (e.g., Thompson, Willemsen, Gooch, & Creem-Regehr, 2004).

4. Discussion

The aim of the present study was to investigate whether eye contact with an avatar in VR would evoke similar attention and emotion related psychophysiological responses as eye contact with a live person. Direct gaze compared to averted gaze evoked larger attention orienting responses (as measured by HR deceleration). The effect of gaze direction on HR deceleration was very small in the VR condition, although not significantly different from the Live condition. However, there was a significant difference between the conditions in the effect of gaze direction on SCRs indexing physiological arousal. Larger SCRs to direct gaze relative to averted gaze were observed only in the Live condition, but not in the VR condition. These results demonstrated that virtual eye contact did not evoke meaningfully large physiological attention orienting or arousal responses. The self-rating data showed instead that subjective arousal was judged as higher in response to direct gaze than averted gaze in both presentation conditions. The questionnaires also indicated that social presence (sense of being in the presence of another person) was reported as higher in the Live condition than in the VR condition.

Previous research has reported larger HR deceleration responses to direct gaze relative to averted gaze both in live interactions and in VR (Akechi et al., 2013; Myllyneva & Hietanen, 2015b; Wieser et al., 2010), but the current study was the first to directly compare the effect in the two presentation conditions. There was no statistically significant difference between the conditions, but the effect was numerically larger in the Live condition than in the VR condition. Analyses showed that if virtual eye contact influences HR deceleration, the effect is quite small (significantly smaller than a medium effect size). Thus, the present results show that, in this kind of a laboratory setting, an avatar's direct gaze does not evoke a meaningfully large effect on HR. Further research with more statistical power is needed to conclude whether the magnitude of the effect indeed differs from the effect in a live interaction. However, we prefer not to give specific recommendations as to how large a sample would be sufficient for answering this question, as sample effect sizes (e.g., $\eta_p^2 = .06$ for the non-significant two-way interaction between Presentation mode and Gaze direction in the present study) are not always an accurate basis for determining sample sizes for subsequent studies (Lakens, 2013).

Another interesting finding was that live faces, as compared to avatar faces, evoked larger HR deceleration responses (averaged over the gaze directions). This could reflect that attention was engaged more by live faces than virtual faces, possibly due to higher social salience and physiological arousal in response to live faces (cf. Vogt, De Houwer, Koster, Van Damme, & Crombez, 2008). While the avatars were quite realistic, they were nevertheless of lower visual fidelity than the live faces. This may have attenuated the experience of social salience and social presence, thereby reducing physiological responses to the avatar. This hypothesis could be tested in the future by varying the visual fidelity of the avatar and investigating whether this modulates physiological responses to the avatar.

It was surprising that virtual eye contact had no effect on SCRs at all. The present VR interaction was designed to evoke the simultaneous experience of seeing another person and the experience of being seen, which have been suggested to be necessary for direct gaze to increase physiological arousal (Myllyneva & Hietanen, 2015a, 2015b). Importantly, however, the VR condition differed from live eye contact in that the visual interaction was technologically mediated: the interaction partners saw each other's avatars rather than real faces. There are two plausible interpretations as to why this could have diminished the physiological arousal response to eye contact. Firstly, the avatars' lower visual fidelity when compared to real faces may have diminished physiological responses to nonverbal cues shown by the avatar. Secondly and perhaps more importantly, participants' awareness that the interaction partner saw one's avatar could have diminished the experience of being the object of another's attention and made it less arousing. Eye contact has been found to increase public self-awareness, such as concern about one's outer appearance (Hietanen et al., 2008; Myllyneva & Hietanen, 2015b). However, this effect might not be triggered in avatar-mediated eye contact, especially when the appearance of one's avatar does not match one's real appearance. Most current VR applications utilize avatars that do not resemble the users, but recently, researchers have reported developments towards creating photorealistic look-alike avatars (i.e., avatars realistically depicting the users themselves; Wei et al., 2019). Future research could investigate whether virtual eye contact involving such avatars would evoke similar physiological and affective responses as live eye contact.

A possible interpretation of the overall findings from the physiological measurements is that virtual eye contact engaged attention similarly to real eye contact, but it did not evoke a similar emotional reaction. However, some caution is warranted when interpreting this pattern of results, as attentional and affective processes are closely interlinked (e.g., Pourtois, Schettino, & Vuilleumier, 2013). HR deceleration and SCRs are typically used as indices of attention orienting and arousal, respectively (e.g., Hietanen et al., 2008; Kolassa & Miltner, 2006), but they are not “pure” representations of these processes. Instead, attention orienting responses are modulated by emotions (e.g., Kreibig, 2010), and conversely, SCRs are related to attention as well as arousal (e.g., Dindo & Fowles, 2008). Thus, our measurements do not allow conclusively disentangling attentional and affective effects.

While the SCRs were not enhanced by direct gaze in VR, the subjective self-ratings showed that participants reported more arousal in response to direct than averted gaze, similarly in both the Live and VR conditions. This finding is in line with evidence showing that direct gaze often influences implicit, automatic affective responses and subjective affective judgments differently (Chen, Helminen, & Hietanen, 2017; Hietanen, 2018). This is because affective judgments are subject to top-down influence from higher-level cognitive and motivational processes, whereas implicit measurements reflect bottom-up automatic affective responses. An avatar’s self-directed gaze may thus be reported as subjectively arousing even if it does not evoke a physiological arousal response. Gaze direction could influence, for instance, the meaning participants give to the interaction, thus altering the subjective experience.

The subjective self-ratings of valence were similar between gaze directions and presentation modes. Previous studies employing self-ratings have resulted in diverse findings regarding the effects of gaze direction on affective valence ratings: some studies have reported higher valence ratings to direct than to averted gaze, some studies have reported the opposite pattern, and the rest have reported no effects at all (for a review, see Hietanen, 2018). Thus, in this respect, the present findings were not surprising. Interestingly, previous research relying on implicit behavioral (e.g., affective priming paradigm) and physiological (e.g., facial electromyography and startle reflex) measures have found direct gaze, relative to other gaze directions, to evoke effects associated with positive affect (Chen, Helminen, & Hietanen, 2017; Chen, Peltola, Dunn, Pajunen, & Hietanen, 2017; Dubey, Ropar, & Hamilton, 2015; Hietanen et al., 2018; Lawson, 2015). To replicate and extend the findings of the present study, future studies investigating the effects of eye contact on affective responses in VR should use also implicit measures sensitive to affective valence in addition to measures sensitive to arousal.

We believe that depicting users' gaze directions with the help of eye trackers could enhance social VR experiences. The present results suggest that virtual eye contact can evoke subjective affective arousal responses, and thus implementing eye gaze information in VR could make virtual interactions more emotionally engaging. Moreover, eye information may facilitate virtual interactions in similar ways as in live interactions. People can spontaneously mentalize what other people see and experience in VR (cf. Gallagher et al., 2002; Kircher et al., 2009), and may therefore draw inferences about others' intentions and mental states based on their gaze directions. Direct gaze can thus be used in natural ways, for instance, to engage another person's attention (cf. Conty et al., 2006), to initiate interactions (cf. Adams & Kleck, 2005), or to convey that there is something important at the gazed-at location (Frischen et al., 2007).

However, the present results also suggest that, at least in current VR applications using non-photorealistic avatars, the applicability of eye gaze may still be somewhat limited. Diminished psychophysiological responses to virtual eye contact (compared to real eye contact) may diminish the effectiveness of eye information, for instance, in increasing users' immersion in VR entertainment, or in making people feel emotionally connected through VR telecommunication. Furthermore, our findings have implications also for therapeutic interventions using VR, such as exposure therapy in treatment of social anxiety (e.g., Anderson et al., 2013) and eye contact practice for children with autism (Elgarf, Abdennadher, & Elshahawy, 2017). A potential benefit of VR exposure is that some patients may be more open to treatment due to experiencing lowered physiological arousal during virtual interactions. However, the lowered arousal levels could somewhat reduce the impact of such treatments, and thus the treatments could benefit from being complemented by more arousing live exposure (cf. Owens & Beidel, 2015). Importantly, all of these social VR applications could possibly be made more potent if future research could identify the conditions that need to be met for eye contact to evoke physiological responses such as enhanced SCRs. If these responses can be reproduced in VR, for instance by using photorealistic look-alike avatars, this could potentially enhance future social VR applications by making them feel more realistic, immersive and engaging.

As a final note, the present findings are relevant for the use of VR in research on social cognitive phenomena. VR has been proposed as a useful tool in this field due to being able to combine the rich interactivity and ecological validity of live experimental designs with the precise experimental control allowed by the use of computers (e.g., Bailenson, Blascovich, Beall, & Loomis, 2003; Pan & Hamilton, 2018). We endorse this idea and agree that VR opens up exciting new possibilities for studying human cognition and behavior. It allows creating experiments that would be difficult, expensive, or unethical to conduct otherwise. At the same time, however, the results of the current experiment highlight that findings from studies conducted in VR do not necessarily apply outside of VR. Just like people respond differently to face pictures and live faces (Hietanen et al., 2008; Pönkänen et al., 2011), there are also important differences between responses to humans encountered in VR and in live situations. Artificial stimuli such as virtual avatars and photographs are extremely useful in psychological research, but perhaps they cannot completely replace live stimuli when studying people’s responses in live social interactions.

References

Adams, R. B. Jr., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion, 5*, 3–11. doi:10.1037/1528-3542.5.1.3

Akechi, H., Senju, A., Uibo, H., Kikuchi, Y., Hasegawa, T., & Hietanen, J. K. (2013). Attention to eye contact in the west and east: Autonomic responses and evaluative ratings. *PLoS One, 8*, e59312. doi:10.1371/journal.pone.0059312

Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E., & Calamaras, M. R. (2013). Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 81*, 751–760. doi:10.1037/a0033559

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. doi:10.1177/0146167203029007002

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Hoboken: Wiley.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*, 49–59. doi:10.1016/0005-7916(94)90063-9

Chen, T., Helminen, T. M., & Hietanen, J. K. (2017). Affect in the eyes: Explicit and implicit evaluations. *Cognition and Emotion, 31*, 1070–1082. doi:10.1080/02699931.2016.1188059

Chen, T., Peltola, M. J., Dunn, R., Pajunen, S. M., & Hietanen, J. K. (2017). Modulation of the eyeblink and cardiac startle reflexes by genuine eye contact. *Psychophysiology, 54*, 1872–1881. doi:10.1111/psyp.12975

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159. doi:10.1037/0033-2909.112.1.155

Conty, L., George, N., & Hietanen, J. K. (2016). Watching eyes effects: When others meet the self. *Consciousness and Cognition, 45*, 184–197. doi:10.1016/j.concog.2016.08.016

Conty, L., Tijus, C., Hugueville, L., Coelho, E., & George, N. (2006). Searching for asymmetries in the detection of gaze contact versus averted gaze under different head views: A behavioural study. *Spatial Vision, 19*, 529-545. doi:10.1163/156856806779194026

Critchley, H. D. (2002). Electrodermal responses: What happens in the brain. *The Neuroscientist, 8*, 132–142. doi:10.1177/107385840200800209

Dawson, M. E., Schell, A. M., & Filion, D. L. (2000). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 200–223). Cambridge, UK: Cambridge University Press.

Dindo, L., & Fowles, D. C. (2008). The skin conductance orienting response to semantic stimuli: Significance can be independent of arousal. *Psychophysiology, 45*, 111-118. doi:10.1111/j.1469-8986.2007.00604.x

Dodds, T. J., Mohler, B. J., & Bülthoff, H. H. (2011). Talk to the virtual hands: Self-animated avatars improve communication in head-mounted display virtual environments. *PloS One, 6*, e25759. doi:10.1371/journal.pone.0025759

Donovan, W. L., & Leavitt, L. A. (1980). Physiologic correlates of direct and averted gaze. *Biological Psychology, 10*, 189-199. doi:10.1016/0301-0511(80)90014-9

Dubey, I., Ropar, D., & Hamilton, A. F. de C. (2015). Measuring the value of social engagement in adults with and without autism. *Molecular Autism, 6*, 35. doi:10.1186/s13229-015-0031-2

Elgarf, M., Abdennadher, S., & Elshahawy, M. (2017). I-interact: A virtual reality serious game for eye contact improvement for children with social impairment. In M. Alcañiz, S. Göbel, M. Ma, M. F. Oliveira, J. Baalsrud Hauge, & T. Marsh (Eds.), *Serious Games. JCSG 2017. Lecture Notes in Computer Science, vol 10622*. doi:10.1007/978-3-319-70111-0_14

Farroni, T., Csibra, G., Simion, F., & Johnson, M., H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, 99*, 9602-9605. doi:10.1073/pnas.152159999

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. doi:10.3758/BF03193146

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133*, 694–724. doi:10.1037/0033-2909.133.4.694

Gale, A., Lucas, B., Nissim, R., & Harpham, B. (1972). Some EEG correlates of face-to-face contact. *British Journal of Social and Clinical Psychology, 11*, 326–332. doi:10.1111/j.2044-8260.1972.tb00822.x

Gale, A., Spratt, G., Chapman, A. J., & Smallbone, A. (1975). EEG correlates of eye contact and interpersonal distance. *Biological Psychology, 3*, 237–245. doi:10.1016/0301-0511(75)90023-X

Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage, 16*, 814–821. doi:10.1006/nimg.2002.1117

Graham, F. K., & Clifton, R. K. (1966). Heart-rate change as a component of the orienting response. *Psychological Bulletin, 65*, 305–320. doi:10.1037/h0023258

Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators and Virtual Environments, 1*, 262–271. doi:10.1162/pres.1992.1.2.262

Helminen, T. M., Kaasinen, S. M., & Hietanen, J. K. (2011). Eye contact and arousal: The effects of stimulus duration. *Biological Psychology, 88*, 124–130. doi:10.1016/j.biopsycho.2011.07.002

Hietanen, J. K. (2018). Affective eye contact: An integrative review. *Frontiers in Psychology, 9*, 1587. doi:10.3389/fpsyg.2018.01587

Hietanen, J. K., Helminen, T. M., Kiilavuori, H., Kylliäinen, A., Lehtonen, H., & Peltola, M. J. (2018). Your attention makes me smile: Direct gaze elicits affiliative facial expressions. *Biological Psychology, 132*, 1–8. doi:10.1016/j.biopsycho.2017.11.001

Hietanen, J. K., Leppänen, J. M., Peltola, M. J., Linna-aho, K., & Ruuhiala, H. J. (2008). Seeing direct and averted gaze activates the approach–avoidance motivational brain systems. *Neuropsychologia, 46*, 2423–2430. doi:10.1016/j.neuropsychologia.2008.02.029

Jarick, M., & Bencic, R. (2019). Eye contact is a two-way street: Arousal is elicited by the sending and receiving of eye gaze information. *Frontiers in Psychology, 10*, 1262. doi:10.3389/fpsyg.2019.01262

Joseph, R. M., Ehrman, K., McNally, R., & Keehn, B. (2008). Affective response to eye contact and face recognition ability in children with ASD. *Journal of the International Neuropsychological Society, 14*, 947-955. doi:10.1017/S1355617708081344

Kampe, K. K., Frith, C. D., & Frith, U. (2003). "Hey John": Signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *The Journal of Neuroscience, 23*, 5258-5263. doi:10.1523/JNEUROSCI.23-12-05258.2003

Kircher, T., Blümel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., . . . Krach, S. (2009). Online mentalising investigated with functional MRI. *Neuroscience Letters, 454*, 176–181. doi:10.1016/j.neulet.2009.03.026

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin, 100*, 78–100. doi:10.1037/0033-2909.100.1.78

Kleinke, C. L., & Pohlen, P. D. (1971). Affective and emotional responses as a function of other person's gaze and cooperativeness in a two-person game. *Journal of Personality and Social Psychology, 17*, 308-313. doi:10.1037/h0030600

Kolassa, I., & Miltner, W. H. (2006). Psychophysiological correlates of face processing in social phobia. *Brain Research, 1118*, 130-141. doi:10.1016/j.brainres.2006.08.019

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology, 84*, 394-421. doi:10.1016/j.biopsycho.2010.03.010

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863. doi:10.3389/fpsyg.2013.00863

Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science, 8*, 355-362. doi:10.1177/1948550617697177

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science, 1*, 259-269. doi:10.1177/2515245918770963

Lawson, R. (2015). I just love the attention: Implicit preference for direct eye contact. *Visual Cognition, 23*, 450–488. doi:10.1080/13506285.2015.1039101

Leavitt, L. A., & Donovan, W. L. (1979). Perceived infant temperament, locus of control, and maternal physiological response to infant gaze. *Journal of Research in Personality, 13*, 267–278. doi:10.1016/0092-6566(79)90018-7

Lyyra, P., Astikainen, P., & Hietanen, J. K. (2018). Look at them and they will notice you: Distractor-independent attentional capture by direct gaze in change blindness. *Visual Cognition, 26*, 25-36. doi:10.1080/13506285.2017.1370052

Lyyra, P., Myllyneva, A., & Hietanen, J. K. (2018). Mentalizing eye contact with a face on a video: Gaze direction does not influence autonomic arousal. *Scandinavian Journal of Psychology, 59*, 360–367. doi:10.1111/sjop.12452

Muthén, L. K., & Muthén, B. O. (1998−2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Myllyneva, A., & Hietanen, J. K. (2015a). The dual nature of eye contact: To see and to be seen. *Social Cognitive and Affective Neuroscience, 11*, 1089–1095. doi:10.1093/scan/nsv075

Myllyneva, A., & Hietanen, J. K. (2015b). There is more to eye contact than meets the eye. *Cognition, 134*, 100–109. doi:10.1016/j.cognition.2014.09.011

Nichols, K., & Champness, B. (1971). Eye gaze and the GSR. *Journal of Experimental Social Psychology, 7*, 623–626. doi:10.1016/0022-1031(71)90024-2

Öhman, A., Hamm, A., & Hugdahl, K. (2000). Cognition and the autonomic system: Orienting, anticipation, and conditioning. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 533–575). Cambridge, UK: Cambridge University Press.

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*, 296–305. doi:10.1007/s10862-014-9454-x

Pan, X., & Hamilton, A. F de C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology, 109*, 395–417. doi:10.1111/bjop.12290

Pejsa, T., Andrist, S., Gleicher, M., & Mutlu, B. (2015). Gaze and attention management for embodied conversational agents. *ACM Transactions on Interactive Intelligent Systems, 5*, 3. doi:10.1145/2724731

Peltola, M. J., Hietanen, J. K., Forssman, L., & Leppänen, J. M. (2013). The emergence and stability of the attentional bias to fearful faces in infancy. *Infancy, 18*, 905–926. doi:10.1111/infa.12013

Plutchik, R. (1980). *Emotion: A Psychoevolutionary Synthesis*. New York, NY : Harper & Row.

Pönkänen, L. M., Alhoniemi, A., Leppänen, J. M., & Hietanen, J. K. (2011). Does it make a difference if I have an eye contact with you or with your picture? An ERP study. *Social Cognitive and Affective Neuroscience, 6*, 486-494. doi:10.1093/scan/nsq068

Pönkänen, L. M., Peltola, M. J., & Hietanen, J. K. (2011). The observer observed: Frontal EEG asymmetry and autonomic responses differentiate between another person's direct and averted gaze when the face is seen live. *International Journal of Psychophysiology, 82*, 180–187. doi:10.1016/j.ijpsycho.2011.08.006

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology, 92*, 492-512. doi:10.1016/j.biopsycho.2012.02.007

Prinsen, J., & Alaerts, K. (2019). Eye contact enhances interpersonal motor resonance: Comparing video stimuli to a live two-person action context. *Social Cognitive and Affective Neuroscience, 14*, 967-976. doi:10.1093/scan/nsz064

Richards, J. E., & Turner, E. D. (2001). Extended visual fixation and distractibility in children from six to twenty-four months of age. *Child Development, 72*, 963–972. doi:10.1111/1467-8624.00328

Schrammel, F., Pannasch, S., Graupner, S., Mojzisch, A., & Velichkovsky, B. M. (2009). Virtual friend or threat? The effects of facial expression and gaze interaction on psychophysiological responses and emotional experience. *Psychophysiology, 46*, 922-931. doi:10.1111/j.1469-8986.2009.00831.x

Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences, 13*, 127–134. doi:10.1016/j.tics.2008.11.009

Sharples, S., Cobb, S., Moody, A., & Wilson, J. R. (2008). Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays, 29*, 58–69. doi:10.1016/j.displa.2007.09.005

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications.* New York, NY: Wiley

Sjouwerman, R., & Lonsdorf, T. (2019). Latency of skin conductance responses across stimulus modalities. *Psychophysiology, 56*, e13307. doi:10.1111/psyp.13307

Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., & Beall, A. C. (2004). Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators and Virtual Environments, 13*, 560–571. doi:10.1162/1054746042545292

Vogt, J., De Houwer, J., Koster, E. H., Van Damme, S., & Crombez, G. (2008). Allocation of spatial attention to emotional stimuli depends upon arousal and not valence. *Emotion, 8*, 880-885. doi:10.1037/a0013981

von Grünau, M., & Anston, C. (1995). The detection of gaze direction: A stare-in-the-crowd effect. *Perception, 24*, 1297-1313. doi:10.1068/p241297

Wei, S., Saragih, J., Simon, T., Harley, A. W., Lombardi, S., Perdoch, M., . . . Sheikh, Y. (2019). VR facial animation via multiview image translation. *ACM Transactions on Graphics, 38*. doi:10.1145/3306346.3323030

Wieser, M. J., Pauli, P., Alpers, G. W., & Mühlberger, A. (2009). Is eye to eye contact really threatening and avoided in social anxiety?—An eye-tracking and psychophysiology study. *Journal of Anxiety Disorders, 23*, 93–103. doi:10.1016/j.janxdis.2008.04.004

Wieser, M. J., Pauli, P., Grosseibl, M., Molzow, I., & Mühlberger, A. (2010). Virtual social interactions in social anxiety—the impact of sex, gaze, and interpersonal distance. *Cyberpsychology, Behavior, and Social Networking, 13*, 547–554. doi:10.1089/cyber.2009.0432

Wilcox, L. M., Allison, R. S., Elfassy, S., & Grelik, C. (2006). Personal space in virtual reality. *ACM Transactions on Applied Perception, 3*, 412–428. doi:10.1145/1190036.1190041

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*, 225–240. doi:10.1162/105474698565686

Notes

[1] We replicated our statistical analyses for the physiological and subjective emotional measurements, including Block order as a between-subjects factor. Block order did not interact with Gaze direction in any of the analyses. Details are provided in supplementary materials.

[2] A limitation in the analyses of the physiological data was that the responses were averaged over several trials. Thus, the analyses could not take into account the within-person variation of the responses within the same condition, which can be significant because attention orienting responses and SCRs tend to habituate quickly after a few trials (Dawson et al., 2000; Öhman et al., 2000). To ensure the results were not affected by aggregating the trials, we replicated the analyses on the trial-level, adjusting for within-person clustering in Mplus version 8 (Bollen & Curran, 2006; Muthén & Muthén, 1998-2017). The results, a latent quadratic curve model for HR (n = 1152) and an intercept-only model for SCR (n = 1020), showed a habituation effect over the trials for SCR but not for HR. Otherwise the results were consistent with the main results of the ANOVAs.

Data availability. The data analyzed in this study are not publicly available to not compromise participant consent, but are available from the corresponding author on reasonable request.