Face Processing in Real and Virtual Faces: An EEG Study

Julija Vaitonytė (j.vaitonyte@tilburguniversity.edu)
Department of Cognitive Science & Artificial Intelligence, Tilburg University, Tilburg, The Netherlands

Maryam Alimardani (m.alimardani@vu.nl)
Department of Computer Science, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Max M. Louwerse (m.m.louwerse@tilburguniversity.edu)
Department of Cognitive Science & Artificial Intelligence, Tilburg University, Tilburg, The Netherlands

Abstract

Previous studies suggested brain differences in the temporal domain when processing real human faces versus virtual agent faces, starting from 400 ms onward. However, few studies directly compared the early and the late face processing stages within one paradigm. Here we conducted an EEG study utilizing real human faces and high-quality virtual agent faces, examining two event-related potentials; the early N170 and the Late Positive Potential (LPP). Results showed identical N170 responses for both face types. However, the LPP response revealed a nuanced distinction, with real human faces evoking a slightly larger LPP compared to virtual agent faces. These results suggest that although virtual agent faces can approach the level of emotional engagement and higher-order evaluation associated with real human faces, human faces remain the most engaging. These findings shed light on the cognitive processes involved in face perception and the potential for intelligent virtual agents in AI and education.

Keywords: face processing; virtual faces; EEG; event-related potential (ERP); N170; LPP

Introduction

Humans have a specialized cognitive and neural machinery designed for face processing (Duchaine & Yovel, 2015), a capability refined over millennia through exposure to the morphological and reflectance patterns of natural human faces (Oh, Dotsch, & Todorov, 2019; Sheehan & Nachman, 2014; Sinha, Balas, Ostrovsky, & Russell, 2006). Recently, these human faces are accompanied by artifacts that replicate the human face appearance with differing levels of fidelity. These include the faces of social robots—physically embodied entities—and virtual agents, whether graphically rendered in 2D or 3D, and designed for verbal and non-verbal interaction with humans (Lugrin, 2021; Vaitonytė, Alimardani, & Louwerse, 2023). Furthermore, the past decade has seen proliferation of synthetic media, images and videos, that depict human faces created using AI-based creation technologies, relying on machine learning techniques, such as Generative Adversarial Networks (GANs) (Karras et al., 2020; Yu et al., 2020) and Diffusion Probabilistic Models (Stypulkowski et al., 2023).

Given the ongoing technological evolution, previous experimental studies examined humans' ability to behaviorally distinguish between real human faces and synthetic ones (Nightingale & Farid, 2022; Vaitonytė, Blomsma, Alimardani, & Louwerse, 2021). While the faces synthesized using GANs are indistinguishable from real faces to human observers (Miller et al., 2023; Nightingale & Farid, 2022; Tucciarelli, Vehar, Chandaria, & Tsakiris, 2022), the faces of virtual agents created using 3D scanning techniques and rendered as 2D images can be distinguished from real human photographs (Vaitonytė et al., 2021). The behavioral studies are augmented by brain imaging studies. Prior work in this field suggests that the brain can, in fact, identify differences between the real and synthetic faces (Moshel, Robinson, Carlson, & Grootswagers, 2022). For instance, Moshel et al. (2022) showed that electroencephalography (EEG) activity can be used to decode whether participants processed real human faces or the faces created using GANs. Different than the behavioral findings that can ascertain whether participants can distinguish between real and synthetic faces, EEG studies allow for monitoring the actual process of determining whether a face is real or synthetic. It may be the case that a human perceiver processes a (real or synthetic) face stimulus through a series of processing stages, with some of these stages indicating differential processing of real vs. synthetic faces. Those brain imaging techniques sensitive in the time domain can help uncover a cascade of different sub-processes regarding face processing.

EEG and magnetoencephalography (MEG) both have a high temporal resolution and thus are particularly suitable to studying face processing as it unfolds over time. Prior MEG work highlighted the dynamic nature of face processing and that it proceeds in a coarse-to-fine fashion (Dobs, Isik, Pantazis, & Kanwisher, 2019; Wardle, Taubert, Teichmann, & Baker, 2020). For example, illusory faces (i.e., objects that resemble faces, also called pareidolia) are processed like faces within the first 100 ms, but within 200 ms the brain disentangled differing representations and represents pareidolia more similar to objects than faces (Dobs et al., 2019; Wardle et al., 2020). Other neural work also suggests temporal unfolding of different face dimensions. For instance, face age and gender are extracted before face identity (Dobs et al., 2019). Specifically, age, gender, and identity are all extracted within the first 100 ms, but the perception of age and gender were shown to arise 20 ms earlier than that of identity. On the other hand, face familiarity appears to be extracted later in time, at around 400 ms and onwards. Dobs et al. (2019) however was not able to clarify whether the familiarity signature in the brain was related to the activation of memories that were linked to a specific familiar individual, an affective

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

response to a familiar face, or a more general familiarity response.

The coarse-to-fine trajectory observed in MEG studies, namely, first processing age and gender, and then identity is also reflected in the EEG studies looking at different event-related brain potentials (ERPs), such as the N170 and the Late Positive Potential (LPP). Previous EEG studies using different types of faces suggested that the earliest phase of face processing is the same for different faces, including real human faces, virtual agent faces, and the faces of dolls (Schindler, Zell, Botsch, & Kissler, 2017; Wheatley, Weinberg, Looser, Moran, & Hajcak, 2011). Specifically, the N170 response, which peaks 140 ms – 200 ms post-stimulus onset, was of comparable magnitude in response to real human faces and doll faces (Wheatley et al., 2011) and in response to real human faces and virtual agent faces (Schindler et al., 2017). On the other hand, a larger LPP response, which peaks 400 – 600 ms post-stimulus onset, was found for real compared to virtual agent faces (Schindler et al., 2017) and for real human faces compared to doll faces (Wheatley et al., 2011). However, there also exists research pointing to a larger LPP for virtual agent than real human faces (Cheetham, Wu, Pauli, & Jancke, 2015). Despite the discrepancies, these results may be interpreted as the evidence that initially the brain is attuned to the identification of the basic face pattern as indicated by the presence of the early N170 response, followed by the processes relating to higher-order evaluation as indicated by the LPP that peaks later in time.

Overall, while there exists some research regarding the neural representation of real human and virtual agent faces, there is scarcity of studies that investigate the processing of real and high-quality virtual agent faces over time. Crucially, there exist inconsistencies in the current literature with respect to the late stage of processing of real human faces and virtual agent faces.

The current paper adds to the existing literature by clarifying whether a stronger LPP response is evoked to real or virtual agent faces, as well as how real human and virtual agent faces are processed dynamically in the brain, a question relevant to different domains within the cognitive sciences, including psychology (face processing), artificial intelligence (embodied agents), linguistics (multimodal communication), and education (intelligent tutoring systems).

We conducted an EEG experiment in which participants were presented with a set of real human faces intermixed with high-quality virtual agent faces. Participants were asked to attend to different faces while their EEG activity was recorded. Based on previous literature (Schindler et al., 2017; Wheatley et al., 2011), we predicted that real human and virtual agent faces would evoke a comparable N170 response, whereas the amplitude of the LPP, associated with elaboration, emotional engagement and episodic memory encoding (Schupp, Flaisch, Stockburger, & Junghöfer, 2006), would differ significantly between real human faces and virtual agent faces.

Method

Participants

Twenty-one students (17 females, 4 males, Age: *Mean* = 19.76, *SD* = 2.14) took part in this experiment and received course credit for their participation. All participants had normal or corrected-to-normal vision and were right-handed. They received information about the experiment and gave informed consent. The experiment was approved by the Research Ethics and Data Management Committee of the Tilburg School of Humanities and Digital Sciences (identification code: REC2019/03).

Stimuli

Images used in the experiment consisted of two types of color images: (1) photographs of human faces, and (2) high-fidelity virtual agent faces (Figure 1). Human face photographs (*N* = 32, half female) were obtained from the Chicago Face Database (CFD) (Ma, Correll, & Wittenbrink, 2015). Since all images came from the CFD, they were all equivalent on low-level image characteristics and aspects such as attractiveness. The images of virtual agent faces (*N* = 32, 18 female faces) were either obtained from the Internet (*N* = 25, 13 female faces) or from the agents that were developed in-house (*N* = 7, 5 female faces) using photogrammetry, a technique of 3D scanning (Foster & Halbstein, 2014). To gather the images from the Internet we used the following criteria: (1) the photographs of the virtual agent face had to be of high quality, (2) the face had to be presented in frontal view, and (3) the face was not covered with hair that obscured facial features. Neither virtual agent, nor human images were manipulated. To prepare stimuli, both human and virtual agent faces were cropped to an oval to expose only the face, removing all non-facial information (e.g., hair). All images had a constant height (800 pixels) and a slightly varying width due to inherent variation in the facial width (from 550 to 650 pixels).

Figure 1: Example of stimuli, showing two high-quality virtual agent faces on the left and two human faces on the right.

Even though our stimuli were not controlled for low-level characteristics, our selected images were comparable in size and viewpoint. We chose not to control for low-level features for two main reasons: (1) even with respect to the early components, such as N170, it has been shown that they are largely unaffected by low-level characteristics of stimuli (Bentin et al., 2007), and (2) the stimuli used in the current study had previously been employed in two behavioral studies exploring perception and memory (Vaitonytė et al., 2021; 2022). Thus, we aimed to maintain consistency in stimuli across studies to facilitate result interpretation.

Procedure

Before commencing the experiment, participants were screened for neurological disorders (e.g., epilepsy and migraine). They then read an information letter, signed an informed consent form, and filled in a demographic questionnaire asking about their age, gender, ethnicity, handedness, and whether they had normal or corrected-to-normal vision. Next, the experimenter provided verbal instructions about the experiment nature, set up and the expectations from the participants (e.g., minimizing movement during EEG recording). The experiment presented face images one by one on the computer screen. Participants were instructed to sit still and focus on each image. No specific cognitive task was assigned following previous literature (Schindler et al., 2017; Wheatley et al., 2011). Once seated, participants first saw a brief introduction and were familiarized with the procedure by being presented with three practice images (one virtual agent face and two real human faces). These practice trials were not included in the analysis. Next, the experimental images were presented. The 32 virtual agent face images and 32 human face images were presented on a white background semi-randomly (different randomization lists were used) using PsychoPy, an open-source software based on Python (Peirce, 2007). The images were displayed on a monitor with a 60 Hz refresh rate and 2560 × 1440 resolution. Each trial consisted of presenting a face stimulus for 1,000 ms, with the interstimulus interval being jittered between 700 ms and 2,700 ms, during which a black fixation cross was presented centrally on the screen. Participants were seated approximately 40 cm from the computer monitor and the images subtended 14.3 × 19.9 degrees of visual angle. The stimulus presentation lasted for about 4 minutes.

EEG Signal Acquisition

Data collection took place in a quiet and dimly lit room. A wireless EEG cap, g.Nautilus Research (g.tec medical engineering GmbH, Austria), was used to record continuous EEG brain signals. This cap is lightweight, with 32 prefixed electrodes over frontal, central, temporal, and parietal regions according to 10-20 International Electrode Placing System (Figure 2). The signal was digitized at 24-bit resolution and recorded at a sampling rate of 250 Hz. The ground electrode was mounted on the right earlobe (there is no reference electrode as the system is bipolar). To acquire high-quality EEG signals, conductive gel was applied at each electrode site to ensure good contact with the scalp and reduce impedance. The continuous EEG data were subjected to an online band-pass filter between 0.5 Hz and 60 Hz, and a notch filter between 48 Hz and 52 Hz.

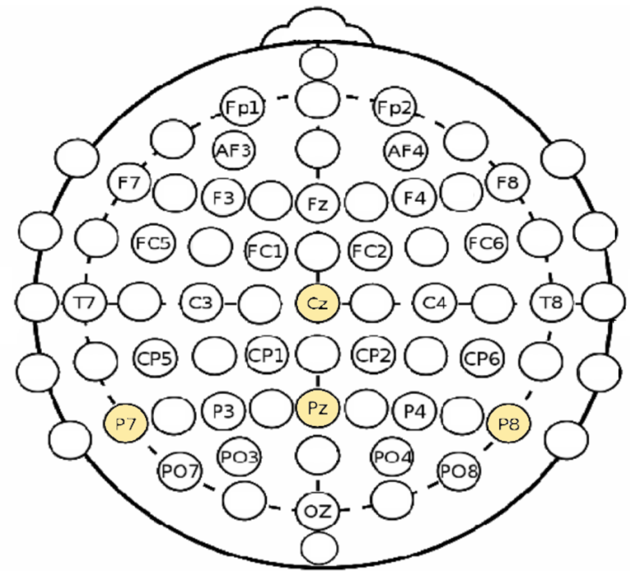

Figure 2: Electrode placement according to 10-20 system. EEG data were acquired using 32 electrodes, of which 4 were used for the analyses, P7, P8, and Cz, Pz as marked in light yellow.

Data Pre-processing and Analysis

EEGLAB (Delorme & Makeig, 2004; Version 2023.0) running under Matlab (V2023a) was used for data pre-processing and analyses. First, data was re-referenced using the Common Average Reference method. Since the data was already filtered online with a low- and high-pass filter, we did not filter the data offline. We then segmented the data into epochs of 1,200 ms, beginning 200 ms before the trial onset (stimulus presentation) and ending at the end of the trial. The 200 ms before trial onset was used for baseline correction. Next, we utilized a semi-automatic rejection routine to eliminate artifacts from the data. Independent Component Analysis (ICA) using the "runica()" function was applied on the epoched data. Additionally, the data were visually inspected to identify and discard any epochs that still contained artifacts following the ICA procedure (a total of 1.56% of trials were rejected for all participants). To calculate the event-related potential, the time-locked average (time-locked to stimulus onset) over all retained trials was computed separately for the two conditions, virtual agent faces (henceforth, agent) and real human faces (henceforth, human), for each participant. Following previous research, N170 was scored as the mean activity between 140 ms and 200 ms from lateral posterior electrodes, P7 and P8, at which N170 is typically maximal (Eimer, 2011). The LPP was scored as the mean activity between 400 ms and 600 ms from midline central and posterior sites: Cz and Pz, where it also tends to be maximal (Hajcak, Moser, & Simons, 2006). The obtained mean amplitude values were subjected to a statistical analysis in R (R Core Team, 2013; version 4.2.2). Since the data were normally distributed paired *t*-tests were used to verify our hypotheses. The signifi-

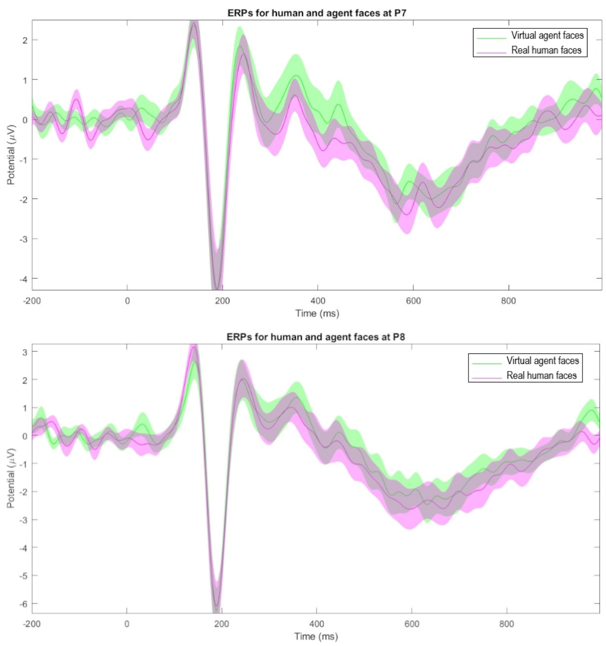

Figure 3: ERPs elicited by agent and human faces at parietal electrode sites, P7 (top) and P8 (bottom). N170 is indicated by a negative deflection at around 180 ms that is identical for both agent and human faces. The shaded area shows the standard error.

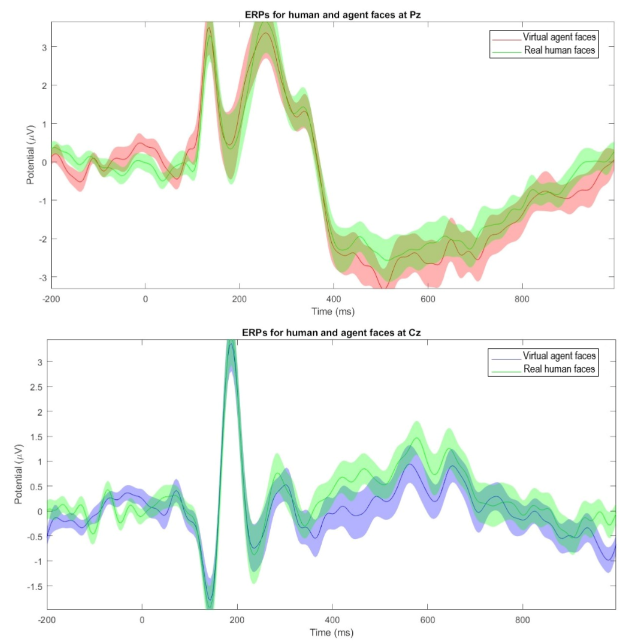

Figure 4: ERPs elicited by agent and human faces at parietal and centro-parietal electrode sites, Pz (top) and Cz (bottom), respectively. LPP tends to be larger for human compared to agent faces as indicated by a more positive deflection at Cz for human than agent faces from 400 ms onwards. The shaded area shows the standard error.

cance threshold was adjusted using the Bonferroni correction method for two comparisons at two different electrode sites ($p < .05/2 = .025$).

Results

As shown in Figure 3, presenting human and agent faces elicited N170 response between 140 ms – 200 ms, which did not differ in amplitude between the two conditions neither at P7, $t(20) = 0.05$, $p = .96$ ($M_{\text{Human}} = -1.34$, $SD_{\text{Human}} = 3.61$; $M_{\text{Agent}} = -1.35$, $SD_{\text{Agent}} = 3.35$), nor at P8, $t(20) = 0.18$, $p = .86$ ($M_{\text{Human}} = -1.92$, $SD_{\text{Human}} = 2.91$; $M_{\text{Agent}} = -1.98$, $SD_{\text{Agent}} = 2.91$).

Presenting human faces elicited LPP responses that showed a trend of being larger than LPP responses to agent faces at Cz, $t(20) = 2.25$, $p = .03$ ($M_{\text{Human}} = 0.84$, $SD_{\text{Human}} = 1.18$; $M_{\text{Agent}} = 0.30$, $SD_{\text{Agent}} = 1.67$) as shown in Figure 4. However, at Pz, the LPP between human and agent faces did not differ, $t(20) = 0.18$, $p = 1.46$ ($M_{\text{Human}} = -2.28$, $SD_{\text{Human}} = 2.10$; $M_{\text{Agent}} = -2.62$, $SD_{\text{Agent}} = 2.22$). Albeit not significant when taking multiple comparisons into account, overall, the results indicated the trend in the predicted direction, with human faces leading to larger late positivity compared to agent faces, while the early N170 response was identical for human and agent conditions.

Discussion

The current study investigated the neural processing time course during the observation of real human faces and high-quality virtual agent faces, in particular, examining the early and the late stages of processing as measured by EEG. The results revealed that both types of faces elicited a comparable N170 response while the LPP response was slightly larger for human faces than virtual agent faces, although the difference between real and virtual agent faces did not reach adjusted significance threshold. Overall, ERPs were in the predicted direction, showing neural processing that was similar for both types of faces at an early stage of observation but then tended to be different at a later stage.

These findings contribute to the existing literature on N170, indicating its responsiveness to various types of faces, including emoticons, stylized faces, doll faces, virtual agent faces, and real human faces (Mustafa & Magnor, 2016; Schindler et al., 2017; Wheatley et al., 2011). Our study demonstrated that high-quality virtual agent faces elicited an N170 response that was indistinguishable from the response evoked by real human faces. This result is particularly interesting in light of our previous behavioral results using largely the same set of virtual agent faces as in the current study (Vaitonytė et al., 2021). Specifically, a computational analysis revealed that virtual agent faces had fewer corneal reflections in the eyes and smoother skin texture compared to real human faces, which was further corroborated in perceptual experiments, showing that observers identified virtual agent faces as agent-like based on these subtle discrepancies in the eyes and the skin. Despite this, as shown by the current study, at the early stage of processing, the brain appears to be less concerned with minute facial details so long as the stimulus presents a face pattern—at least when assessed by N170.

Regarding the LPP component, our findings largely align with prior work by Schindler et al. (2017), showing larger LPP values for human faces relative to virtual agent faces. However, the similarity of the LPP waveforms in the current study for high-quality virtual agent and real human faces suggests that these virtual agent faces might partially tap into higher-order evaluation, affective processes, and memory encoding expected for fully-fledged human faces (Schindler et al., 2017; Wheatley et al., 2011).

Does this then mean that the brain cannot differentiate between the highest quality computer-generated/synthetic faces and real human faces? While a recent study by Moshel et al. (2022) found that it was possible for machine learning classifiers to decode from EEG activity whether GAN images of faces were perceived as realistic or not, behavioral studies show that perceptually humans struggle to tell apart GAN faces from real human faces (Miller et al., 2023; Nightingale & Farid, 2022; Tucciarelli et al., 2022). The current neural results similarly point to a dissociation between behavior and the brain. Specifically, while observers can distinguish real human faces from virtual agent faces (Vaitonytė et al., 2021), the differentiation between these two classes of faces only begins to manifest at a later stage of neural processing.

In the future, it would be advantageous to investigate and compare the neural processing of real human faces and virtual agent faces based on different late ERP components, including the LPP and N400. The N400 component, found to be indicative of integrating and processing semantic information (Kutas & Federmeier, 2011), has been associated with the uncanny valley response in humans—a phenomenon where artificial entities, such as virtual agent faces, provoke discomfort in humans as they become increasingly human-like (Mustafa, Guthe, Tauscher, Goesele, & Magnor, 2017); see Vaitonytė et al. (2023) for an alternative perspective on N400 and the uncanny valley.

The current findings however need to be considered in a broader context, acknowledging their limitations. First, as with many experimental studies, our participant population primarily included young adult female students. Follow-up experiments ought to employ a more varied participant sample to understand how these results generalize to individuals of differing age and gender. Furthermore, age and gender taken into account, interactions between the individual characteristics of the human participants and the stimuli need to be taken into account. Finally, a more informative approach to examining the processing of real and virtual agent faces in terms of both temporal and spatial scales would be to involve a simultaneous EEG-fMIR recording, as demonstrated in Liu, Huang, McGinnis-Deweese, Keil, and Ding (2012). This hybrid neuroimaging method would make it possible to understand how LPP amplitudes vary with the engagement of specific regions within the visual cortex as well as deep subcortical structures, such as the amygdala, known to be involved in emotional processing of stimuli (Liu et al., 2012; Sabatinelli, Keil, Frank, & Lang, 2013).

In summary, our results suggest that face processing evolves from a broad to specific pattern, where the perceptual system initially exhibits broad responsiveness to different faces, but at a later stage, it shows a heightened response to faces of high perceptual quality.

Conclusion

Current EEG results provide evidence that high-quality virtual agent faces elicit an early-stage neural response, represented by the N170 component, that is indistinguishable from the response evoked by real human faces. This suggests that the early stage of face processing is primarily concerned with the presence of a face template rather than minute details or perceptual quality, at least when assessed by N170. Additionally, our findings suggest that virtual agent faces can to some extent tap into the processes associated with fully-fledged faces; however, a slight difference in the LPP response between real human faces and virtual agent faces suggests that real human faces may still evoke stronger associations with higher-order evaluation. These results contribute to the understanding of the temporal unfolding of face processing and highlight the potential of high-quality virtual agent faces to engage similar neural mechanisms as real human faces. The relevance of these findings spans across the development of embodied agents, their utilization in fields ranging from healthcare to education, and adds to our understanding of the psychological processes underlying human face perception.

References

Bentin, S., Taylor, M. J., Rousselet, G. A., Itier, R. J., Caldara, R., Schyns, P. G., ... Rossion, B. (2007). Controlling interstimulus perceptual variance does not abolish n170 face sensitivity. *Nature Neuroscience*, *10*(7), 801–802. doi: 10.1038/nn0707-801

Cheetham, M., Wu, L., Pauli, P., & Jancke, L. (2015). Arousal, valence, and the uncanny valley: Psychophysiological and self-report findings. *Frontiers in Psychology*, *6*, 981. doi: 10.3389/fpsyg.2015.00981

Delorme, A., & Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Dobs, K., Isik, L., Pantazis, D., & Kanwisher, N. (2019). How face perception unfolds over time. *Nature Communications*, *10*(1), 1258. doi: 10.1038/s41467-019-09239-1

Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, *1*, 393–416. doi: 10.1146/annurev-vision-082114-035518

Eimer, M. (2011). The face-sensitive n170 component of the event-related brain potential. In M. J. A. Calder G. Rhodes

& J. Haxby (Eds.), *Oxford handbook of face perception.* Oxford University Press.

Foster, S., & Halbstein, D. (2014). *Integrating 3d modeling, photogrammetry and design*. Springer.

Hajcak, G., Moser, J. S., & Simons, R. F. (2006). Attending to affect: appraisal strategies modulate the electrocortical response to arousing pictures. *Emotion*, *6*(3), 517–522. doi: 10.1037/1528-3542.6.3.517

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of stylegan. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 8110–8119).

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the n400 component of the event-related brain potential (erp). *Annual Review of Psychology*, *62*, 621–647. doi: 10.1146/annurev.psych.093008.131123

Liu, Y., Huang, H., McGinnis-Deweese, M., Keil, A., & Ding, M. (2012). Neural substrate of the late positive potential in emotional processing. *Journal of Neuroscience*, *32*(42), 14563–14572. doi: 10.1523/JNEUROSCI.3109-12.2012

Lugrin, B. (2021). Introduction to socially interactive agents. In *The handbook on socially interactive agents: 20 years of research on embodied conversational agents, intelligent virtual agents, and social robotics volume 1: Methods, behavior, cognition* (pp. 1–20). doi: 10.1145/3477322.3477324

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*, 1122–1135. doi: 10.3758/s13428-014-0532-5

Miller, E. J., Steward, B. A., Witkower, Z., Sutherland, C. A., Krumhuber, E. G., & Dawel, A. (2023). Ai hyperrealism: Why ai faces are perceived as more real than human ones. *Psychological Science*, *34*(12), 1390–1403. doi: 10.1177/09567976231207095

Moshel, M. L., Robinson, A. K., Carlson, T. A., & Grootswagers, T. (2022). Are you for real? decoding realistic ai-generated faces from neural activity. *Vision Research*, *199*, 108079. doi: 10.1016/j.visres.2022.108079

Mustafa, M., Guthe, S., Tauscher, J.-P., Goesele, M., & Magnor, M. (2017). How human am i? eeg-based evaluation of virtual characters. In *Proceedings of the 2017 chi conference on human factors in computing systems* (pp. 5098–5108). doi: 10.1145/3025453.3026043

Mustafa, M., & Magnor, M. (2016). Eeg based analysis of the perception of computer-generated faces. In *Proceedings of the 13th european conference on visual media production (cvmp 2016)* (pp. 1–10). doi: 10.1145/2998559.2998563

Nightingale, S. J., & Farid, H. (2022). Ai-synthesized faces are indistinguishable from real faces and more trustworthy. *Proceedings of the National Academy of Sciences*, *119*(8), e2120481119. doi: 10.1073/pnas.2120481119

Oh, D., Dotsch, R., & Todorov, A. (2019). Contributions of shape and reflectance information to social judgments from faces. *Vision Research*, *165*, 131–142. doi: 10.1016/j.visres.2019.10.010

Peirce, J. W. (2007). Psychopy—psychophysics software in python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13. doi: 10.1016/j.jneumeth. 2006.11.017

R Core Team. (2013). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Sabatinelli, D., Keil, A., Frank, D. W., & Lang, P. J. (2013). Emotional perception: correspondence of early and late event-related potentials with cortical and subcortical functional mri. *Biological Psychology*, *92*(3), 513–519. doi: 10.1016/j.biopsycho.2012.04.005

Schindler, S., Zell, E., Botsch, M., & Kissler, J. (2017). Differential effects of face-realism and emotion on event-related brain potentials and their implications for the uncanny valley theory. *Scientific Reports*, *7*(1), 1–13. doi: 10.1038/srep45003

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Progress in Brain Research*, *156*, 31–51. doi: 10.1016/S0079-6123(06)56002-9

Sheehan, M. J., & Nachman, M. W. (2014). Morphological and population genomic evidence that human faces have evolved to signal individual identity. *Nature Communications*, *5*(1), 1–10. doi: 10.1038/ncomms5800

Sinha, P., Balas, B., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE*, *94*(11), 1948–1962. doi: 10.1109/JPROC.2006.884093

Stypulkowski, M., Vougioukas, K., He, S., Zieba, M., Petridis, S., & Pantic, M. (2023). *Diffused heads: Diffusion models beat gans on talking-face generation. corr abs/2301.03396 (2023).*

Tucciarelli, R., Vehar, N., Chandaria, S., & Tsakiris, M. (2022). On the realness of people who do not exist: The social processing of artificial faces. *Iscience*, *25*(12). doi: 0.1016/j.isci.2022.105441

Vaitonytė, J., Alimardani, M., & Louwerse, M. M. (2022). Corneal reflections and skin contrast yield better memory of human and virtual faces. *Cognitive Research: Principles and Implications*, *7*(1), 94. doi: 10.1186/s41235-022-00445-y

Vaitonytė, J., Alimardani, M., & Louwerse, M. M. (2023). Scoping review of the neural evidence on the uncanny valley. *Computers in Human Behavior Reports*, *9*, 100263. doi: 10.1016/j.chbr.2022.100263

Vaitonytė, J., Blomsma, P. A., Alimardani, M., & Louwerse, M. M. (2021). Realism of the face lies in skin and eyes: Evidence from virtual and human agents. *Computers in Human Behavior Reports*, *3*, 100065. doi: 10.1016/j.chbr.2021.100065

Wardle, S. G., Taubert, J., Teichmann, L., & Baker, C. I.

(2020). Rapid and dynamic processing of face pareidolia in the human brain. *Nature Communications*, *11*(1), 4518. doi: 10.1038/s41467-020-18325-8

Wheatley, T., Weinberg, A., Looser, C., Moran, T., & Hajcak, G. (2011). Mind perception: Real but not artificial faces sustain neural activity beyond the n170/vpp. *PloS one*, *6*(3), e17960. doi: 10.1371/journal.pone.0017960

Yu, J., Xu, X., Gao, F., Shi, S., Wang, M., Tao, D., & Huang, Q. (2020). Toward realistic face photo–sketch synthesis via composition-aided gans. *IEEE Transactions on Cybernetics*, *51*(9), 4350–4362. doi: 10.1109/TCYB.2020.2972944