



Emotional Valence Recognition on Virtual, Robotic, and Human Faces: a Comparative Study

Lisa Cerda, Pierluigi Graziani, Jonathan Del-Monte

► To cite this version:

Lisa Cerda, Pierluigi Graziani, Jonathan Del-Monte. Emotional Valence Recognition on Virtual, Robotic, and Human Faces: a Comparative Study. *Journal of Technology in Behavioral Science*, 2021, 6 (1), pp.151 - 158. <10.1007/s41347-020-00172-5>. <hal-03171138>

HAL Id: hal-03171138

<https://hal.science/hal-03171138v1>

Submitted on 16 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Emotional Valence Recognition on Virtual, Robotic, and Human Faces: a Comparative Study

Lisa Cerda, Pierluigi Graziani, Jonathan Del-Monte

► To cite this version:

Lisa Cerda, Pierluigi Graziani, Jonathan Del-Monte. Emotional Valence Recognition on Virtual, Robotic, and Human Faces: a Comparative Study. Journal of Technology in Behavioral Science, 2020, 6, pp.151 - 158. 10.1007/s41347-020-00172-5 . hal-03171138

HAL Id: hal-03171138

<https://hal.archives-ouvertes.fr/hal-03171138>

Submitted on 16 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Emotional Valence Recognition on Virtual, Robotic, and Human Faces: a Comparative Study

Lisa Cerda^{1,2} · Pierluigi Graziani^{1,2} · Jonathan Del-Monte^{1,2}

Received: 2 May 2020 / Revised: 25 September 2020 / Accepted: 1 October 2020

Abstract

With the advent of new technologies in everyday life, leisure, or therapy, we will increasingly interact with a non-human virtual character. Understanding facial expressions and intentions of these virtual agents is important to enable them to achieve their goals. The objective of our study is to assess whether expressions are perceived as being positive or negative on faces more or less similar to those of humans. Eighty-three undergraduate students took part in a computerized emotion recognition task. The participants had to identify whether each face expressed a positive or a negative emotion. Eight different faces (human, avatar, mesh, and robot) were shown 38 times each on a computer screen. Each face was represented by a photo. Response time and the number of correct responses were recorded. Our research has raised important points: the accuracy and time taken for emotion recognition were found to be similar on human or avatar faces. On the other hand, as soon as these faces were too ambiguous or schematic, emotion recognition capacities were found to be diminished.

Keywords Avatar · Emotion recognition · Facial expression

Introduction

Humans communicate in sophisticated ways through many channels such as sound, vision, and even touch. This is how they convey important information. The spoken language occupies such an important place that we sometimes forget the importance of non-verbal signs. However, facial expressions play a major role in social interactions (Adolphs 1999). They make it possible to convey messages and information essential for survival (i.e., fear) or more broadly for communication allowing empathy (Frith 2009). The ability to recognize and understand facial expressions is therefore essential for social relationships. This is a universal skill for social interaction

(Ekman 1993). The processing of emotional expressions involves several subcortical regions such as the amygdala and occipital and temporal cortex regions (Dubal et al. 2011). In general, humans recognize positive facial expressions faster than negative expressions because they do not require the analysis of the whole face (i.e., just a smiling mouth can signify joy) (Leppänen and Hietanen 2004). With the advent of new technologies, such as virtual reality, augmented reality, and artificial intelligence, we can wonder whether humans are capable of the same performance in recognizing emotions in robots and avatars. Artificial agents become an integral part of human life in leisure, the professional world or in therapy. We are therefore required to interact daily with virtual non-human characters or those personified in the form of a physical robot (Churamani et al. 2016).

Some studies have shown that virtual characters are capable of provoking social reactions equivalent to those experienced before humans (Garau et al. 2008). Expressive robots are considered as more friendly, more human and allow engagement among human users (Eyssel et al. 2010; Hall et al. 2014). Electromyographic studies have shown that photographs of robots and humans opening and closing their respective hands induce similar automatic imitation responses (Oberman et al. 2007; Oztot et al. 2005). But what about the emotions expressed by robots and other non-human agents?

✉ Lisa Cerda
lisa.cerda30@gmail.com

Pierluigi Graziani
pierluigi.graziani@unimes.fr

Jonathan Del-Monte
jonathan.del-monte@unimes.fr

¹ Social Psychology Laboratory, 29, avenue Robert-Schuman, 13621 Aix-en-Provence, cedex 1, France

² Nîmes University, 5 Avenue du Dr Georges Salan, 30 000 Nîmes, France

It has been shown in a study that the emotions expressed by a virtual face are recognized comparably to the emotions expressed by natural facial expressions (Spencer-Smith et al. 2001). There seem to be minor differences in the observation time of the mouth area. Users observed virtual agents for a longer time period (Joyal et al. 2014). Moreover, natural and virtual facial expressions caused comparable brain activation (Moser et al. 2007; Hortensius et al. 2018). However, unlike natural faces, the emotions expressed by avatars were mainly confused with neutrality. This characteristic indicates that the emotions expressed by the avatars were not confused with completely different emotions; when people were unsure of the emotion expressed, there was a tendency to judge the expression as being neutral. (Dyck et al. 2008). It would thus seem that the emotions conveyed by avatars are not always clear. However, we know that ambiguity will be listed as a negative valence on avatars (Cheetham et al. 2015). Nevertheless, avatars are capable of evoking a sense of social presence (Biocca et al. 2003) which highlights our ability to identify their emotions. It has been shown that virtual characters transmit social information but are also perceived as social agents exercising a social influence on observers (Bailenson et al. 2003; Pertaub et al. 2001). Naive users (e.g., children) consider emotional feedback of artificial agents like real emotions (Kahn Jr. et al. 2012). Users can read social clues like emotions in virtual faces (Jack and Schyns 2017; Todorov et al. 2015).

However, the link between avatar social influence and avatar graphics seems complex. Indeed, with a high level of realism, the avatar perception can lead to an unpleasant impression for an observer. This phenomenon is described as uncanny valley (Mori 1970). When the observer perceives a non-humanoid robot, he pays attention to the few elements similar to humanoids. The feelings of familiarity then increase in the observer. Although when the robot looks very similar to a humanoid, the observer pays attention to the few non-humanoid elements. There then appears a feeling of strangeness towards the robot. The robot is not seen as an agent awkwardly imitating a human but as a human incapable of acting like his fellows. Nonetheless, beyond this valley, the robot is accepted as a humanoid similar to the observer. The human visual system accepts all types of faces whether human, artificial, robotics, virtual, drawn, schematic, or geometric (Seyama and Nagayama 2007). It still remains necessary to adapt the physical appearance of robots and other virtual agents, taking into account the emotional and psychological impact it can have on humans. It is therefore necessary to find a balance between a schematic, anthropomorphic, and ultra-realistic face to allow communication and social cooperation in the various fields of use of virtual agents (i.e., training, leisure, therapy). Hence, the aim of our study is to make an update about our emotion recognition skills and to place a slider on the scale of realism to understand whether the human clue is necessary or deleterious for emotion recognition.

The objective of our research is to assess whether emotional valences are perceived in the same way on different types of faces (human, avatar, mesh and robot).

Materials and Methods

Participants

Eighty-three undergraduate students of Nîmes University participated for course credits. The sample size was determined by statistical requirements and the number of volunteers in the first year of psychology at the University of Nîmes. The sample was predominantly female (87.95%), young ($M = 20.31$, $SD = 4.74$), and educated (mean years of education = 12.07, $SD = 0.376$). Candidates suffering from psychotic disorders, substance abuse, major visual disturbance, or epilepsy were excluded. The totality of the data was provided by the participants. For all subjects, the level of depressive symptoms was assessed with the Beck Depression Inventory-III (BDI-II; Beck et al. 1996) and social anxiety using the Liebowitz-Social-Anxiety-Scale (LSAS; Liebowitz 1987). A complete description of the study was made to the participants after which written, informed consent was obtained.

Material

State-Trait Anxiety Inventory (STAI; (Spielberger 1983))

The STAI is a widely used instrument for measuring the subject's general trait anxiety and current state anxiety. The valid French translation of the 20-item current anxiety state instrument (Spielberger 1983) was administered before the experiment to control a potential interpretation bias. A high level of anxiety is associated with a more pessimistic interpretation of an ambiguous stimulus (Lazzeri et al. 2015). The STAI score was controlled for; scores under 46 were considered as weak.

Emotion Recognition Task

Upon the administration of the instructions, a learning task consisting of 10 faces was conducted. This learning time was attributed to familiarize participants with the keyboard; the goal was to limit incorrect responses due to a lack of familiarization with the material. The experimental task comprises of eight different faces (Fig. 1) that were presented thirty-eight times (a total of 304 faces) on a computer screen (60 Hz, 439 mm, 128 DPI). Viewing was unrestrained at a distance of approximately 600 mm. The faces were categorized as follows: human, avatar, robot, and mesh (a frame of avatar represented by the base of avatar face without texture,

115
116
117

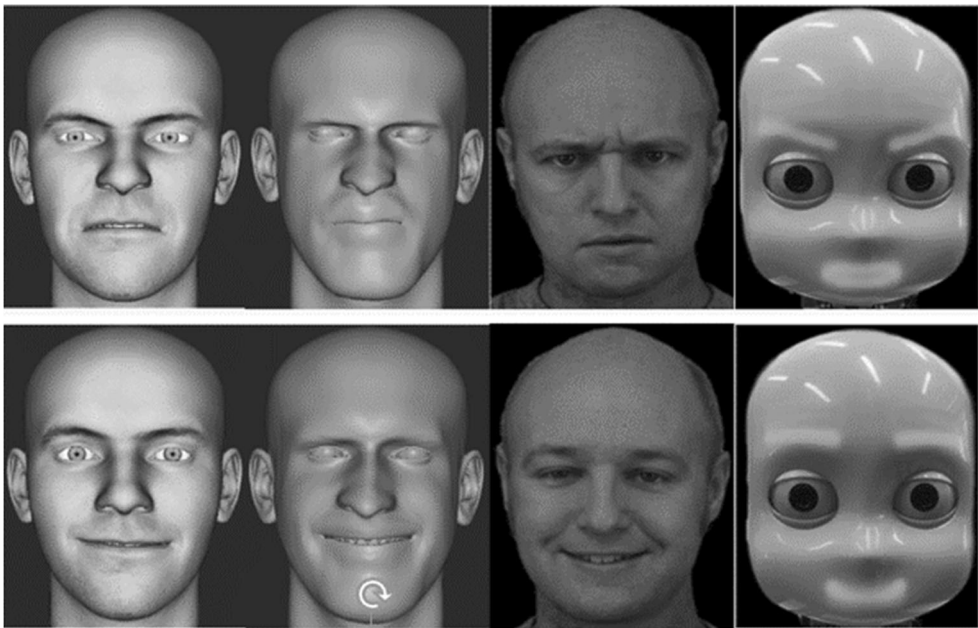
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134

135
136
137

138
139
140
141
142
143
144
145

146
147Q4
148
149
150
151
152
153
154
155
156
157

Fig. 1 On the top row: negative valence stimuli (Avatar, Mesh, Human, Robot); on the bottom row: positive valence stimuli (Avatar, Mesh, Human, Robot)



Q5 158 i.e., skin color, eye details). Each face mimics a facial expres- 188
159 sion having a positive and then a negative valence (Fig. 1). All 189
160 faces were 2D photos in front of a black background. Lighting 190
161 was similar for each face and there were no sound effects. 191
162 Each participant was required to choose whether the facial 192
163 expression represented a positive or a negative valence, via 193
164 the keyboard. Keyboard letters were hidden, only letters S 194
165 and L were showed. “S” was to be used for “positive” re- 195
166 sponse and “L” for “negative” response. Correct responses 196
167 and individual response times were recorded. 197

168 **Procedure**

169 The participants in the present study were volunteers. Each 198
170 participant signed an informed consent form and completed a 199
171 general questionnaire about gender, age, native language, vi- 200
172 sual disturbances, academic degrees, family background, vid- 201
173 eo game habits, virtual reality experiences, and psychiatric 202
174 background. Current state anxiety and depression levels were 203
175 measured using STAI and BDI. Participants then completed 204
176 the preliminary questionnaire after which they were seated in 205
177 front of a computer screen. The keyboard was customized as 206
178 per the requirements of the experimental task. The experi- 207
179 menter verbally explained the instructions: “You will see dif- 208
180 ferent faces on the screen. For each face, you have to choose 209
181 whether the facial expression is positive or negative. You have 210
182 to push the corresponding button on the keyboard. Your re-
183 sponse time will be recorded. This is not a rapidity exercise,
184 but please respond as spontaneously as possible with the least
185 possible errors.” After this learning task, participants started
186 the main task. A face appeared on the screen, and participants
187 had to identify the valence of the emotion represented. To

make a definitive choice, participants had to press the key
corresponding to the identified valence. After each choice,
another face appeared on the screen. The experiment was con-
ducted in an office at the Nîmes University; the participants
were allowed to leave upon completion of the experimental
task.

Statistical Analyses

The Friedman test was used to compare the type of face or
valence for the entire sample. The Wilcoxon test was
employed for 2 by 2 comparison. The level of significance
was set to $p < .05$. Response time is provided in milliseconds.
Correct and incorrect responses are given as a percentage.

Results

Demographic and Clinical Measures

Eighty-three undergraduate students of Nîmes University par-
ticipated. The sample was predominantly female (87.95%),
young ($M = 20.31$, $SD = 4.74$), and educated (mean years of
education = 12.07, $SD = 0.376$). Low scores on depression
($M_{BDI} = 13.17$, $SD = 10.06$) and low current state anxiety
levels were reported. ($M_{STAI} = 37$, $SD = 11.97$).

Results According to the Type of Face

In this study, “response time” corresponds to the time spent to
select a response on the keyboard, and “correct response”

corresponds to the correct classification of the valence of the face by the participants (Fig. 2).

Means According to the Type of Face

Friedman test on response time (Table 1) and percentage of correct response score ($M_{\text{human}} = 94.95$; $SD_{\text{human}} = 7.74$; $M_{\text{avatar}} = 95.26$; $SD_{\text{avatar}} = 6.46$; $M_{\text{mesh}} = 91.28$; $SD_{\text{mesh}} = 13.81$; $M_{\text{robot}} = 92.93$; $SD_{\text{robot}} = 11.43$) reveal a difference between the type of face for response time ($M = 8208.815$, $SD = 93.60$; $\chi^2_r = 17.501$, $p = 0.001$).

To be more precise, participants took lesser time to differentiate the valence of the emotion of human faces compared with the other types of faces. There was no difference between other types of the face except for robots; they were associated to longer response times for the recognition of emotional valence.

Results According to the Valence of Expression

Comparison of the Type of Face According to Emotional Valence

Comparison means showed that participants needed more time to determine negative valence of human ($Z = -3.201$, $p = 0.001$, $d = 0.322$) and robot faces ($Z = -2.470$, $p = 0.014$, $d = 0.251$). However, they need lesser time to determine negative valence of mesh faces ($Z = -2.302$, $p = 0.021$, $d = 0.266$) (Fig. 3). Regarding the percentage of correct responses (Fig. 2), there is a significant difference for mesh. Participants recognized positive valence expressions lesser ($Z = -4.577$, $p = 0.000$, $d = 0.622$).

Comparison According to the Type of Face for each Valence

Friedman test on response time and percentage of correct response scores revealed a difference between the types of the face for response time ($M = 1014.93$, $SD = 253.00$; $\chi^2_r =$

10.475 , $p = 0.015$) and the percentage of correct responses ($M = 91.91$, $SD = 17.91$; $\chi^2_r = 14.066$, $p = 0.003$) for positive valence expression, and only for the percentage of correct responses ($M = 95.30$, $SD = 9.35$; $\chi^2_r = 14.798$, $p = 0.002$) for negative valence expression.

Positive Valence For positive valence expressions, the participants took lesser time to differentiate the emotional valence of the human face compared with mesh and robot type of face (Table 1). There was no difference with the avatar. Participants needed more time to respond for mesh face in comparison with the avatar, but they needed equivalent time in comparison to the robot. Regarding the percentage of correct responses in positive valence expression, participants made more mistakes for mesh face in comparison with the other types of faces ($Z_{\text{human/mesh}} = -3.537$, $p = 0.000$, $d = 0.457$; $Z_{\text{avatar/mesh}} = -3.621$, $p = 0.000$, $d = 0.508$; $Z_{\text{mesh/robot}} = -2.258$, $p = 0.024$, $d = 0.316$). For other faces, participants had equivalent performance ($Z_{\text{human/avatar}} = -0.210$, $p = 0.834$, $d = 0.0488$; $Z_{\text{human/robot}} = -1.661$, $p = 0.097$, $d = 0.152$; $Z_{\text{avatar/robot}} = -1.551$, $p = 0.121$, $d = 0.205$).

Negative Valence For the percentage of correct responses in the negative valence expression, participants made more mistakes for mesh face in comparison with the other types of faces ($Z_{\text{human/mesh}} = -3.250$, $p = 0.001$, $d = 0.219$; $Z_{\text{avatar/mesh}} = -3.018$, $p = 0.003$, $d = 0.239$; $Z_{\text{mesh/robot}} = -3.155$, $p = 0.002$, $d = 0.360$). For other faces, participants had an equivalent performance ($Z_{\text{human/avatar}} = -0.815$, $p = 0.415$, $d = 0.010$; $Z_{\text{human/robot}} = -0.577$, $p = 0.564$, $d = 0.189$; $Z_{\text{avatar/robot}} = -0.570$, $p = 0.569$, $d = 0.185$) (Fig. 4).

Discussion

The results of our study comply with literature. Indeed, our results show the fastest response for human faces in comparison with other faces. Moreover, the results show the fastest

Fig. 2 Percentage of correct and incorrect responses according to valence face

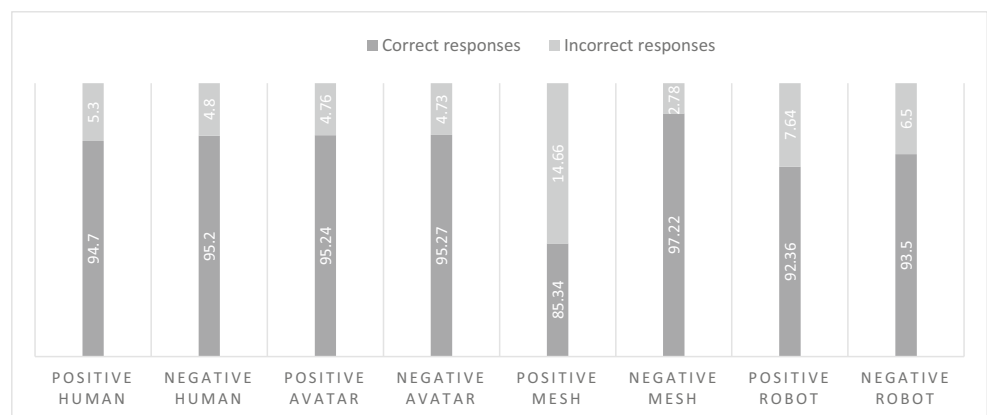


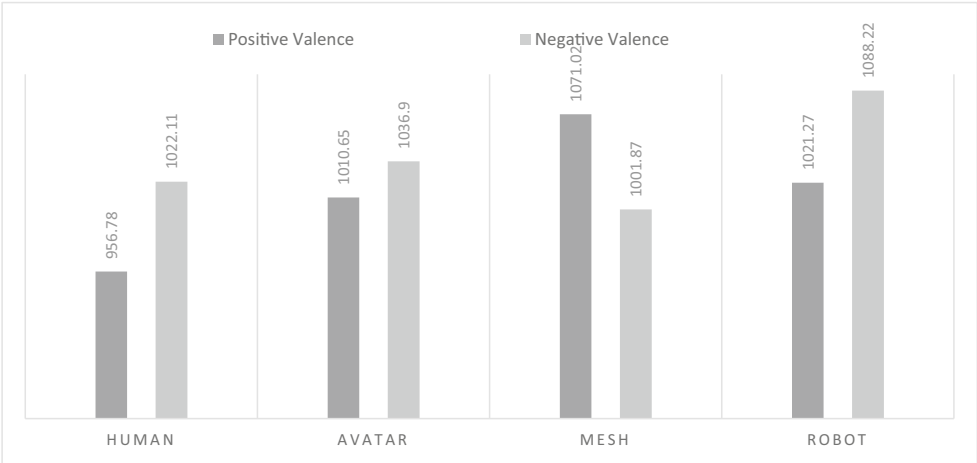
Table 1 Comparison of means according to the type of face for response time

		Means	SD	Statistic (Z)	p value	Effect size
	Sample					
	Human/avatar	1978.89/2047.55	369.70/446.08	-2.048	0.041*	0.167
	Human/mesh	197.889/2072.89	369.70/453.169	-2.888	0.004*	0.227
	Human/robot	1978.89/2109.49	369.70/472.022	-3.814	0.000*	0.308
	Avatar/mesh	2047.55/2072.89	446.08/453.169	-1.171	0.241	0.057
	Avatar/robot	2047.55/2109.49	446.08/472.022	-2.361	0.018*	0.135
	Mesh/robot	2072.89/2109.49	453.169/472.022	-0.445	0.656	0.079
	Positive valence condition					
	Human/avatar	956.78/1010.65	184.11/239.28	-1.966	0.049*	0.252
	Human/mesh	956.78/1071.02	184.11/301.68	-4.032	0.000*	0.457
	Human/robot	956.78/1021.27	184.11/263.72	-2.029	0.042*	0.283
	Avatar/mesh	1010.65/1071.02	239.28/301.68	-2.888	0.004*	0.221
	Avatar/robot	1010.65/1021.27	239.28/263.72	-0.867	0.386	0.042
	Mesh/robot	1071.02/1021.27	301.68/263.72	-1.898	0.058	0.175

recognition of positive facial expressions in comparison with negative facial expressions. According to Leppänen and Hietanen (2004), humans recognize positive facial expressions faster than negative facial expressions because positive expressions are characterized by obvious and unique characteristics such as the smile (Adolphs et al. 2002). Negative expression requires an in-depth analysis of facial traits. Another interesting finding was that the recognition of expression among avatars did not differ from the recognition of expressions among human faces. Similar results have been shown in previous studies. Expressions of the virtual face are as well recognized as those of natural faces (Spencer-Smith et al. 2001). Moreover, both activate the amygdala in the brain (Dyck et al. 2008). These results reflect habituation of human traits which facilitates their understanding. This could explain why we find opposite results with less “human” faces. Indeed, our study shows slower responses for avatars

compared with responses given for human faces. More particularly, the discrimination of positive expressions requires significantly more time for robot faces than human faces. Even if humans tend to attribute human traits to robots (Breazeal 2003; Wendt and Berg 2009), other studies show that these faces are still perceived differently from human faces. It all depends on the human characteristics of the robot’s face; the more anthropomorphic it is, the easier it will be to recognize its expressions (Raffard et al. 2016). We can, therefore, imagine that the more robots share common features with us, the more natural the communication, which would be a good sign for the acceptance of social robots (Wendt and Berg 2009). Studies using iClub’s face, like us, have found similar results: participants respond more quickly to negative valence expressions than to positive valence expressions (Pais et al. 2013). These results are consistent with a recent meta-analysis conducted by Nummenmaa and Calvo (2015), which showed that

Fig. 3 Means of response time for valence for each type of face



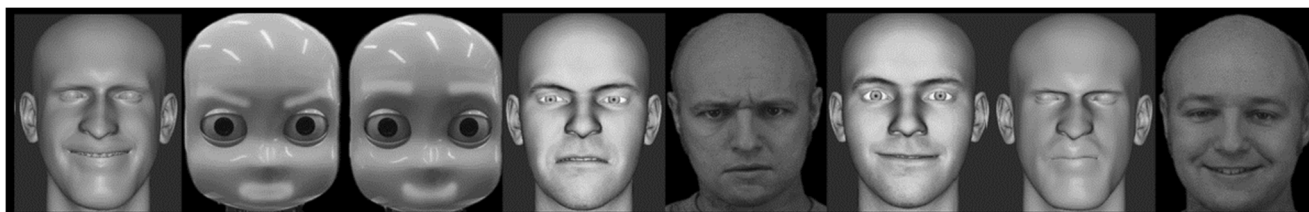


Fig. 4 Ranking of faces from worst to best performance in recognition of the valence of emotion (positive mesh, negative robot, positive robot, negative avatar, negative human, positive avatar, negative mesh, positive human)

happy faces are better recognized in photographs, whereas angry/sad faces are recognized better in schematic faces. This notion seems to be the explanatory key of the results obtained for the “mesh” face. Indeed, negative emotions were identified faster than positive emotions. Some studies have shown that when a face is ambiguous it is attributed to a more negative valence than when it is less ambiguous (Cheetham et al. 2015; Lazzeri et al. 2015). This may be why we obtained a lesser number of correct responses for this face type. Participants interpret it more negatively and attribute a negative valence even for expressions with a positive valence. Authors explain this tendency by our innate predisposition to treat strangers with caution (Zajonc 1998). Moreover, the skin reflects our identity, health care, and other individual characteristics (Alkawaz and Basori 2012). This is what makes us human, and warm. For example, the lack of wrinkles affects emotion perception and recognition (Milcent et al. 2019). We can assume that the mesh stimulus does not allow the collection of relevant information. It would then be considered as an ambiguous stimulus treated negatively by the participants. The realism needed to ward off the ambiguity of facial expressions may lead to the uncanny valley issue. There comes a time when if the agent looks too similar while retaining certain flaws, these appear to be particularly monstrous and disturbing, as when observing a zombie, for example. The degree of familiarity contrasts too starkly with the flaws that remind us that the agent is not a living being when human behavior is expected from him. However, beyond this valley, familiarity reaches a second peak, and humanoids are accepted again. This theory, although controversial, has had a huge impact on robotics and 3D animation. Designers are advised not to attempt to reach the second familiarity peak present after the uncanny valley given the risk of not reaching it (MacDorman 2019) It is because they do not claim to reach the second peak that the avatars presented in our study were able to achieve such good results. Identifying them as a humanoid agent was enough to understand them and interpret their facial expressions. This maintained a relevant degree of affinity and familiarity. The use of this type of face seems relevant for social skills training or every training with non-human agents. Individuals recognize and understand these agents as well as human faces. Moreover, clinical populations with social cognition deficit show altered recognition skill of

avatars’ emotions (Dyck et al. 2010). That makes non-human agents very good training tools, like avatars (Joyal et al. 2014). Their use makes it possible to have training environments very close to the natural environments known to the patient. Despite poorer performance for mesh faces, these faces allow better recognition of negative emotions. Their poor global score is due to difficulties in identifying positive emotions. In contrast, negative emotions are much better and much faster recognized than on other types of faces. This type of face is not to be banned and could be useful for other therapeutic purposes. For example, this type of face can be used to increase the difficulty in an exposure exercise for patients with social anxiety. Once again, this is proof that there is no point in pushing realism to extremes in virtual reality to achieve interesting environmental and social immersion.

Limitations

Our study still has some limitations. We would have liked more males to participate in order to balance the sample. It has been repeatedly proven (see Lawrence et al. 2015) that women tend to identify emotions better than men. A higher proportion of men would have made it possible to obtain more generalizable results for the general population. In addition, recruiting participants among first-year students made our sample young. However, a difference in abilities has been shown in the identification of emotions according to age. For example, Mill et al. (2009) have shown that older individuals recognize negative emotions (sadness and anger) less accurately than younger individuals. Another limitation is that our study involves photos. By definition, images are static, and certain emotions may be more difficult to identify in this manner. We know, for example, that in robots, certain emotions (anger, disgust, fear) are better recognized when they are physically present compared with when displayed on a screen (Lazzeri et al. 2015). To overcome this absence of physical reality, the same virtual reality assessment could be relevant to have a situation closer to a real physical encounter. Thus, the sense of presence and the sense of social presence generated by the different types of faces could be measured, informing us of their understanding by the subjects.

With respect to the clinical implications and therapeutic objectives, it may be interesting to test these stimuli on a

clinical population in order to observe if the results are similar so as to adapt the therapeutic tools.

Conclusion

Our research has raised important points: Emotion recognition requires similar processing time or human and avatar faces. On the other hand, as soon as these faces are too ambiguous or schematic, our capacities are diminished. Theories about the uncanny valley also explain that too much realism of virtual or robotic faces is a guarantee of unpleasant impression. It therefore seems essential to take these results into account while using virtual therapies, for example. A realistic avatar that evokes an unpleasant sensation may not have the desired effect. It is therefore advisable to prefer the use of an avatar but one which retains features of virtual characters in order to promote a sense of presence essential to digital therapies, as well as to avoid the generation of negative sensations. In virtual reality, graphic realism is not a sine qua non condition for the sense of presence for which, in technical terms, material comfort and a sense of familiarity are enough (Sjölie 2012).

Acknowledgments Thanks to Vrutti Joshi, psychologist, for her proof-reading and help for writing process.

Funding We are grateful to the Provence-Alpes-Côte d'Azur Regional Council for his financial support.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, 3(12), 11.

Adolphs, R., Baron-Cohen, S., & Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. *Journal of Cognitive Neuroscience*, 14(8), 1264–1274. <https://doi.org/10.1162/089892902760807258>.

Alkawaz, M. H., & Basori, A. H. (2012). The effect of emotional colour on creating realistic expression of avatar. Proceedings of the 11th ACM SIGGRAPH international conference on virtual-reality continuum and its applications in industry - VRCAI '12, 143. <https://doi.org/10.1145/2407516.2407555>.

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, 29(7), 819–833. <https://doi.org/10.1177/0146167203029007002>.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck depression inventories-IA and-II in psychiatric outpatients. *Journal of Personality Assessment*, 67(3), 588–597. https://doi.org/10.1207/s15327752jpa6703_13.

Bioeca, F., Harms, C., & Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: Review and suggested

criteria. In *Presence: Teleoperators and Virtual Environments*, 456–480.

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1–2), 119–155. [https://doi.org/10.1016/S1071-5819\(03\)00018-1](https://doi.org/10.1016/S1071-5819(03)00018-1).

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

Churamani, N., Cruz, F., Griffiths, S., & Barros, P. (2016). iCub: Learning emotion expressions using human reward.

Dubal, S., Foucher, A., Jouvent, R., & Nadel, J. (2011). Human brain spots emotion in non humanoid robots. *Social Cognitive and Affective Neuroscience*, 6(1), 90–97. <https://doi.org/10.1093/scan/nsq019>.

Dyck, M., Winbeck, M., Leiberg, S., Chen, Y., Gur, R. C., & Mathiak, K. (2008). Recognition profile of emotions in natural and virtual faces. *PLoS One*, 3(11), e3628. <https://doi.org/10.1371/journal.pone.0003628>.

Dyck, M., Winbeck, M., Leiberg, S., Chen, Y., & Mathiak, K. (2010). Virtual faces as a tool to study emotion recognition deficits in schizophrenia. *Psychiatry Research*, 179(3), 247–252. <https://doi.org/10.1016/j.psychres.2009.11.004>.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, 48(4), 384–392. <https://doi.org/10.1037/0003-066X.48.4.384>.

Eyssel, F., Hegel, F., Horstmann, G., & Wagner, C. (2010). Anthropomorphic inferences from emotional nonverbal cues: A case study. *19th International Symposium in Robot and Human Interactive Communication*, 646–651. <https://doi.org/10.1109/ROMAN.2010.5598687>.

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3453–3458. <https://doi.org/10.1098/rstb.2009.0142>.

Hall, J., Tritten, T., Rowe, A., Pipe, A., Melhuish, C., & Leonards, U. (2014). Perception of own and robot engagement in human–robot interactions and their dependence on robotics knowledge. *Robotics and Autonomous Systems*, 62(3), 392–399. <https://doi.org/10.1016/j.robot.2013.09.012>.

Hortensius, R., Hekele, F., & Cross, E. S. (2018). The perception of emotion in artificial agents. *IEEE Transactions on Cognitive and Developmental Systems*, 10(4), 852–864. <https://doi.org/10.1109/TCDS.2018.2826921>.

Jack, R. E., & Schyns, P. G. (2017). Toward a social psychophysics of face communication. *Annual Review of Psychology*, 68(1), 269–297. <https://doi.org/10.1146/annurev-psych-010416-044242>.

Joyal, C. C., Jacob, L., Cigna, M.-H., Guay, J.-P., & Renaud, P. (2014). Virtual faces expressing emotions: An initial concomitant and construct validity study. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00787>.

Kahn Jr., P. H., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., Ruckert, J. H., & Shen, S. (2012). “Robovie, you’ll have to go into the closet now”: Children’s social and moral relationships with a humanoid robot. *Developmental Psychology*, 48(2), 303–314. <https://doi.org/10.1037/a0027033>.

Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00761>.

Lazzeri, N., Mazzei, D., Greco, A., Rotesi, A., Lanatà, A., & De Rossi, D. E. (2015). Can a humanoid face be expressive? A psychophysiological investigation. *Frontiers in Bioengineering and Biotechnology*, 3. <https://doi.org/10.3389/fbioe.2015.00064>.

Leppänen, J. M., & Hietanen, J. K. (2004). Positive facial expressions are recognized faster than negative facial expressions, but why?

507	<i>Psychological Research Psychologische Forschung</i> , 69(1–2), 22–	American Psychological Association. https://doi.org/10.1037/	544
508	29. https://doi.org/10.1007/s00426-003-0157-2 .	e705412011-025.	545
509	Liebowitz, M. R. (1987). Social phobia. <i>Modern Problems of</i>	Raffard, S., Bortolon, C., Khoramshahi, M., Salesse, R. N., Burca, M.,	546
510	<i>Pharmacopsychiatry</i> , 22, 141–173. https://doi.org/10.1159/	Marin, L., Bardy, B. G., Billard, A., Macioce, V., & Capdevielle, D.	547
511	000414022.	(2016). Humanoid robots versus humans: How is emotional valence	548
512	MacDorman, K. F. (2019). La Vallée de l'Étrange de Mori Masahiro:	of facial expressions recognized by individuals with schizophrenia?	549
513	Importance et impact sur l'esthétique et la conception des robots. <i>e-</i>	An exploratory study. <i>Schizophrenia Research</i> , 176(2), 506–513.	550
514	<i>Phaistos</i> , VII-2. https://doi.org/10.4000/ephaistos.5333 .	https://doi.org/10.1016/j.schres.2016.06.001 .	551
515	Milcent, A.-S., Geslin, E., Kadri, A., & Richir, S. (2019). Expressive	Seyama, J., & Nagayama, R. (2007). The uncanny valley: Effect of realism	552
516	virtual human: Impact of expressive wrinkles and pupillary size on	on the impression of artificial human faces. <i>Presence</i> , 16, 337–351.	553
517	emotion recognition. <i>Proceedings of the 19th ACM international</i>	https://doi.org/10.1162/pres.16.4.337 .	554
518	<i>conference on intelligent virtual agents</i> , 215–217. https://doi.org/	Sjölie, D. (2012). Presence and general principles of brain function.	555
519	10.1145/3308532.3329446.	<i>Interacting with Computers</i> , 24(4 SI), 193–202.	556
520	Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences	Spencer-Smith, J., Wild, H., Innes-Ker, A. H., Townsend, J., Duffy, C.,	557
521	in emotion recognition ability: A cross-sectional study. <i>Emotion</i> ,	Edwards, C., Ervin, K., Merritt, N., & Pair, J. W. (2001). Making	558
522	9(5), 619–630. https://doi.org/10.1037/a0016562 .	faces: Creating three-dimensional parameterized models of facial	559
523	Mori, M. (1970). Bukimi no tani [the uncanny valley]. <i>Energy</i> , 7, 33–35.	expression. <i>Behavior Research Methods, Instruments, &</i>	560
524	Moser, E., Derntl, B., Robinson, S., Fink, B., Gur, R. C., & Grammer, K.	<i>Computers</i> , 33(2), 115–123. https://doi.org/10.3758/BF03195356 .	561
525	(2007). Amygdala activation at 3T in response to human and avatar	Spielberger, C. D. (1983). <i>Manual for the state-trait anxiety inventory</i>	562
526	facial expressions of emotions. <i>Journal of Neuroscience Methods</i> ,	<i>STAI (Form Y) (« Self-Evaluation Questionnaire »)</i> . http://ubir.	563
527	161(1), 126–133. https://doi.org/10.1016/j.jneumeth.2006.10.016 .	buffalo.edu/xmlui/handle/10477/1873	564
528	Nummenmaa, L., & Calvo, M. G. (2015). Dissociation between recogni-	Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015).	565
529	tion and detection advantage for facial expressions: A meta-analysis.	Social attributions from faces: Determinants, consequences, accuracy,	566
530	<i>Emotion</i> , 15(2), 243–256. https://doi.org/10.1037/emo0000042 .	and functional significance. <i>Annual Review of Psychology</i> ,	567
531	Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to	66(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-	568
532	face: Blocking facial mimicry can selectively impair recognition of	143831.	569
533	emotional expressions. <i>Social Neuroscience</i> , 2(3–4), 167–178.	Wendt, C. S., & Berg, G. (2009). Nonverbal humor as a new dimension	570
534	https://doi.org/10.1080/17470910701391943 .	of HRI. <i>RO-MAN 2009 - The 18th IEEE International Symposium</i>	571
535	Oztop, E., Franklin, D. W., Chaminade, T., & Cheng, G. (2005). Human–	<i>on Robot and Human Interactive Communication</i> , 183–188. https://	572
536	humanoid interaction: Is a humanoid robot perceived as a human?	doi.org/10.1109/ROMAN.2009.5326230 .	573
537	<i>International Journal of Humanoid Robotics</i> , 537–559.	Zajonc, R. B. (1998). Emotions. In <i>The handbook of social psychology</i> ,	574
538	Pais, A. L., Argall, B. D., & Billard, A. G. (2013). Assessing interaction	<i>Vols. 1–2, 4th ed</i> (p. 591–632). McGraw-Hill.	575
539	dynamics in the context of robot programming by demonstration.		
540	<i>International Journal of Social Robotics</i> , 5(4), 477–490. https://doi.	Publisher's Note Springer Nature remains neutral with regard to jurisdic-	576
541	org/10.1007/s12369-013-0204-0 .	tional claims in published maps and institutional affiliations.	577
542	Pertaub, D.-P., Slater, M., & Barker, C. (2001). An experiment on fear of		
543	public speaking in virtual reality: (705412011-025) [data set].		
544			
545			
546			
547			
548			
549			
550			
551			
552			
553			
554			
555			
556			
557			
558			
559			
560			
561			
562			
563			
564			
565			
566			
567			
568			
569			
570			
571			
572			
573			
574			
575			

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES.

- Q1. Please check if the affiliations are presented correctly.
- Q2. Please check if the section headings are assigned to appropriate levels.
- Q3. Ref. "Garau et al., 2008" is cited in the body but its bibliographic information is missing. Kindly provide its bibliographic information in the list.
- Q4. Please check if the edits to the sentence starting "Upon the administration of the instructions..." retained the intended meaning. Otherwise, kindly amend.
- Q5. Please check if the edits to the sentence starting "Each face mimics a facial expression having..." retained the intended meaning. Otherwise, kindly amend.
- Q6. A citation for Fig. 2 was inserted here so that figure citations are in sequential order. Please check if appropriate. Otherwise, insert citation(s) in sequential order.
- Q7. Table 1 contains entries in boldface but without significance. If deemed to have significance, please provide their significance in the form of a table footnote; otherwise, please set them upright.
- Q8. Missing citation for Figure 4 was inserted here. Please check if appropriate. Otherwise, please provide citation for Figure 4. Note that the order of main citations of figures/tables in the text must be sequential.
- Q9. Please check if the edits to the sentence starting "The use of this type of face seems relevant..." retained the intended meaning. Otherwise, kindly amend.