



**HAL**  
open science

# 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-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, 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<sup>1,2</sup> · Pierluigi Graziani<sup>1,2</sup> · Jonathan Del-Monte<sup>1,2</sup>

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; Oztop 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

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

<sup>2</sup> Nîmes University, 5 Avenue du Dr Georges Salan, 30 000 Nîmes, France

1  
2  
5  
6  
7  
8  
9  
10  
11

23  
Q2 24

25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35

Q1

36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53 Q3  
54  
55  
56  
57  
58  
59  
60  
61

62 It has been shown in a study that the emotions expressed by  
63 a virtual face are recognized comparably to the emotions  
64 expressed by natural facial expressions (Spencer-Smith et al.  
65 2001). There seem to be minor differences in the observation  
66 time of the mouth area. Users observed virtual agents for a  
67 longer time period (Joyal et al. 2014). Moreover, natural and  
68 virtual facial expressions caused comparable brain activation  
69 (Moser et al. 2007; Hortensius et al. 2018). However, unlike  
70 natural faces, the emotions expressed by avatars were mainly  
71 confused with neutrality. This characteristic indicates that the  
72 emotions expressed by the avatars were not confused with  
73 completely different emotions; when people were unsure of  
74 the emotion expressed, there was a tendency to judge the ex-  
75 pression as being neutral. (Dyck et al. 2008). It would thus seem  
76 that the emotions conveyed by avatars are not always clear.  
77 However, we know that ambiguity will be listed as a negative  
78 valence on avatars (Cheetham et al. 2015). Nevertheless, avatars  
79 are capable of evoking a sense of social presence (Biocca  
80 et al. 2003) which highlights our ability to identify their emotions.  
81 It has been shown that virtual characters transmit social  
82 information but are also perceived as social agents exercising a  
83 social influence on observers (Bailenson et al. 2003; Pertaub  
84 et al. 2001). Naive users (e.g., children) consider emotional  
85 feedback of artificial agents like real emotions (Kahn Jr. et al.  
86 2012). Users can read social clues like emotions in virtual faces  
87 (Jack and Schyns 2017; Todorov et al. 2015).

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

The objective of our research is to assess whether emotion-  
al 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

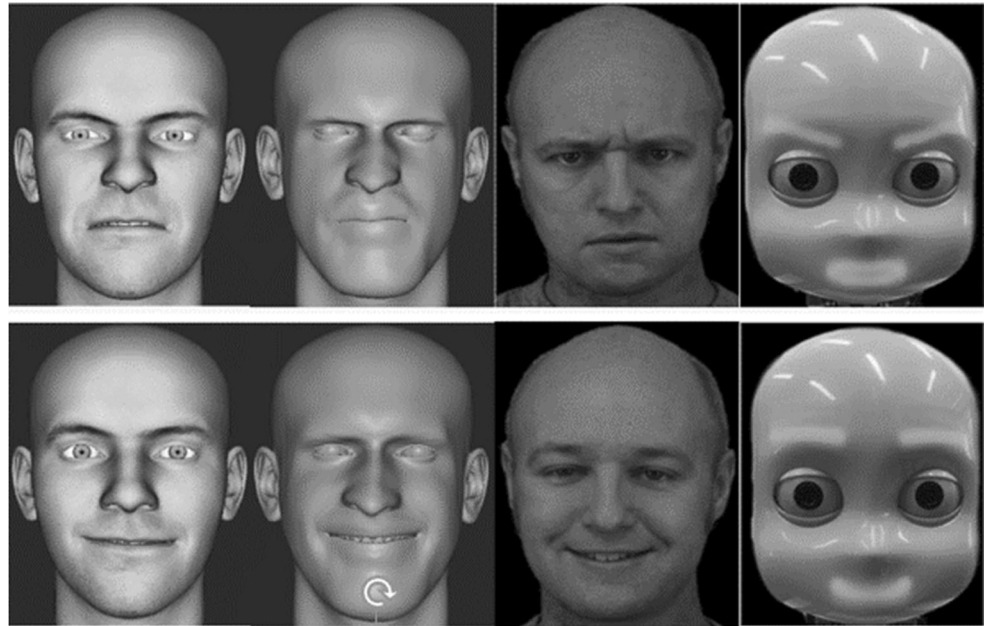
#### State-Trait Anxiety Inventory (STAI; (Spielberger 1983))

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,

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



158 i.e., skin color, eye details). Each face mimics a facial expression having a positive and then a negative valence (Fig. 1). All  
 159 faces were 2D photos in front of a black background. Lighting was similar for each face and there were no sound effects.  
 160 Each participant was required to choose whether the facial expression represented a positive or a negative valence, via  
 161 the keyboard. Keyboard letters were hidden, only letters S and L were showed. “S” was to be used for “positive” response  
 162 and “L” for “negative” response. Correct responses and individual response times were recorded.

**Procedure**

168  
 169 The participants in the present study were volunteers. Each participant signed an informed consent form and completed a  
 170 general questionnaire about gender, age, native language, visual disturbances, academic degrees, family background, video  
 171 game habits, virtual reality experiences, and psychiatric background. Current state anxiety and depression levels were  
 172 measured using STAI and BDI. Participants then completed the preliminary questionnaire after which they were seated in  
 173 front of a computer screen. The keyboard was customized as per the requirements of the experimental task. The experimenter  
 174 verbally explained the instructions: “You will see different faces on the screen. For each face, you have to choose  
 175 whether the facial expression is positive or negative. You have to push the corresponding button on the keyboard. Your  
 176 response time will be recorded. This is not a rapidity exercise, but please respond as spontaneously as possible with the least  
 177 possible errors.” After this learning task, participants started the main task. A face appeared on the screen, and participants  
 178 had to identify the valence of the emotion represented. To

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

**Statistical Analyses**

194  
 195 The Friedman test was used to compare the type of face or  
 196 valence for the entire sample. The Wilcoxon test was  
 197 employed for 2 by 2 comparison. The level of significance  
 198 was set to  $p < .05$ . Response time is provided in milliseconds.  
 199 Correct and incorrect responses are given as a percentage.

**Results**

**Demographic and Clinical Measures**

200  
 201 Eighty-three undergraduate students of Nîmes University participated.  
 202 The sample was predominantly female (87.95%),  
 203 young ( $M = 20.31$ ,  $SD = 4.74$ ), and educated (mean years of  
 204 education = 12.07,  $SD = 0.376$ ). Low scores on depression  
 205 ( $M_{BDI} = 13.17$ ,  $SD = 10.06$ ) and low current state anxiety  
 206 levels were reported. ( $M_{STAI} = 37$ ,  $SD = 11.97$ ).  
 207

**Results According to the Type of Face**

208  
 209 In this study, “response time” corresponds to the time spent to  
 210 select a response on the keyboard, and “correct response”

211 corresponds to the correct classification of the valence of the  
 Q6 212 face by the participants (Fig. 2).

213 **Means According to the Type of Face**

214 Friedman test on response time (Table 1) and percentage of  
 215 correct response score ( $M_{\text{human}} = 94.95$ ;  $SD_{\text{human}} = 7.74$ /  
 216  $M_{\text{avatar}} = 95.26$ ;  $SD_{\text{avatar}} = 6.46$ ;  $M_{\text{mesh}} = 91.28$ ;  $SD_{\text{mesh}} =$   
 217  $13.81$ ;  $M_{\text{robot}} = 92.93$ ;  $SD_{\text{robot}} = 11.43$ ) reveal a difference be-  
 218 tween the type of face for response time ( $M = 8208.815$ ,  $SD =$   
 219  $93.60$ ;  $\chi^2_r = 17.501$ ,  $p = 0.001$ ).

220 To be more precise, participants took lesser time to differ-  
 221 entiate the valence of the emotion of human faces compared  
 222 with the other types of faces. There was no difference between  
 223 other types of the face except for robots; they were associated  
 224 to longer response times for the recognition of emotional  
 225 valence.

226 **Results According to the Valence of Expression**

227 **Comparison of the Type of Face According to Emotional**  
 228 **Valence**

229 Comparison means showed that participants needed more time  
 230 to determine negative valence of human ( $Z = -3.201$ ,  $p = 0.001$ ,  
 231  $d = 0.322$ ) and robot faces ( $Z = -2.470$ ,  $p = 0.014$ ,  $d = 0.251$ ).  
 232 However, they need lesser time to determine negative valence of  
 233 mesh faces ( $Z = -2.302$ ,  $p = 0.021$ ,  $d = 0.266$ ) (Fig. 3).  
 234 Regarding the percentage of correct responses (Fig. 2), there is  
 235 a significant difference for mesh. Participants recognized posi-  
 236 tive valence expressions lesser ( $Z = -4.577$ ,  $p = 0.000$ ,  $d =$   
 237  $0.622$ ).

238 **Comparison According to the Type of Face for each Valence**

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

10.475,  $p = 0.015$ ) and the percentage of correct responses 242  
 (M = 91.91, SD = 17.91;  $\chi^2_r = 14.066$ ,  $p = 0.003$ ) for positive 243  
 valence expression, and only for the percentage of correct 244  
 responses ( $M = 95.30$ ,  $SD = 9.35$ ;  $\chi^2_r = 14.798$ ,  $p = 0.002$ ) 245  
 for negative valence expression. 246

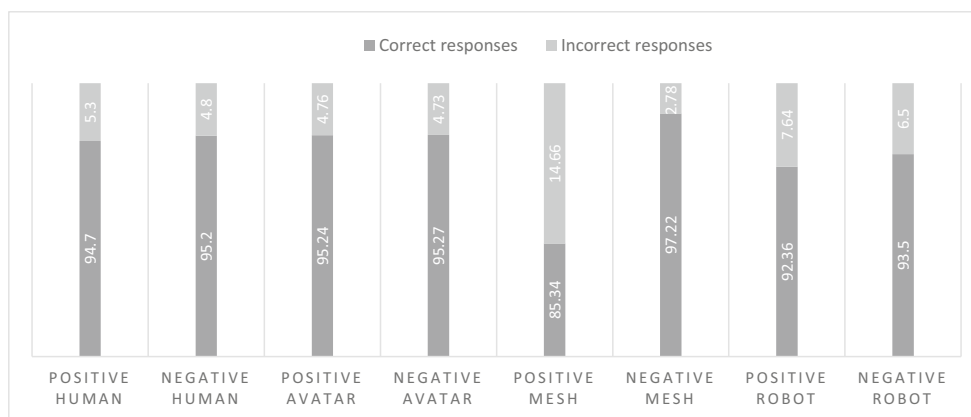
**Positive Valence** For positive valence expressions, the partic- 247  
 248 ipants took lesser time to differentiate the emotional valence  
 of the human face compared with mesh and robot type of face 249  
 (Table 1). There was no difference with the avatar. 250  
 Participants needed more time to respond for mesh face in 251  
 comparison with the avatar, but they needed equivalent time 252  
 in comparison to the robot. Regarding the percentage of cor- 253  
 rect responses in positive valence expression, participants 254  
 made more mistakes for mesh face in comparison with the 255  
 other types of faces ( $Z_{\text{human/mesh}} = -3.537$ ,  $p = 0.000$ ,  $d =$  256  
 $0.457$ / $Z_{\text{avatar/mesh}} = -3.621$ ,  $p = 0.000$ ,  $d = 0.508$ / $Z_{\text{mesh/robot}} =$  257  
 $-2.258$ ,  $p = 0.024$ ,  $d = 0.316$ ). For other faces, participants 258  
 had equivalent performance ( $Z_{\text{human/avatar}} = -0.210$ ,  $p =$  259  
 $0.834$ ,  $d = 0.0488$ / $Z_{\text{human/robot}} = -1.661$ ,  $p = 0.097$ ,  $d = 0.152$ /  
 $Z_{\text{avatar/robot}} = -1.551$ ,  $p = 0.121$ ,  $d = 0.205$ ). 260  
 261

**Negative Valence** For the percentage of correct responses in 262  
 263 the negative valence expression, participants made more mis-  
 takes for mesh face in comparison with the other types of faces 264  
 ( $Z_{\text{human/mesh}} = -3.250$ ,  $p = 0.001$ ,  $d = 0.219$ / $Z_{\text{avatar/mesh}} = -$  265  
 $3.018$ ,  $p = 0.003$ ,  $d = 0.239$ / $Z_{\text{mesh/robot}} = -3.155$ ,  $p = 0.002$ , 266  
 $d = 0.360$ ). For other faces, participants had an equivalent per- 267  
 formance ( $Z_{\text{human/avatar}} = -0.815$ ,  $p = 0.415$ ,  $d = 0.010$ / $Z_{\text{human/}}$  268  
 $\text{robot} = -0.577$ ,  $p = 0.564$ ,  $d = 0.189$ / $Z_{\text{avatar/robot}} = -0.570$ ,  $p =$  269  
 $0.569$ ,  $d = 0.185$ ) (Fig. 4). 270Q8

271 **Discussion**

272 The results of our study comply with literature. Indeed, our 273  
 274 results show the fastest response for human faces in compar-  
 ison with other faces. Moreover, the results show the fastest

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



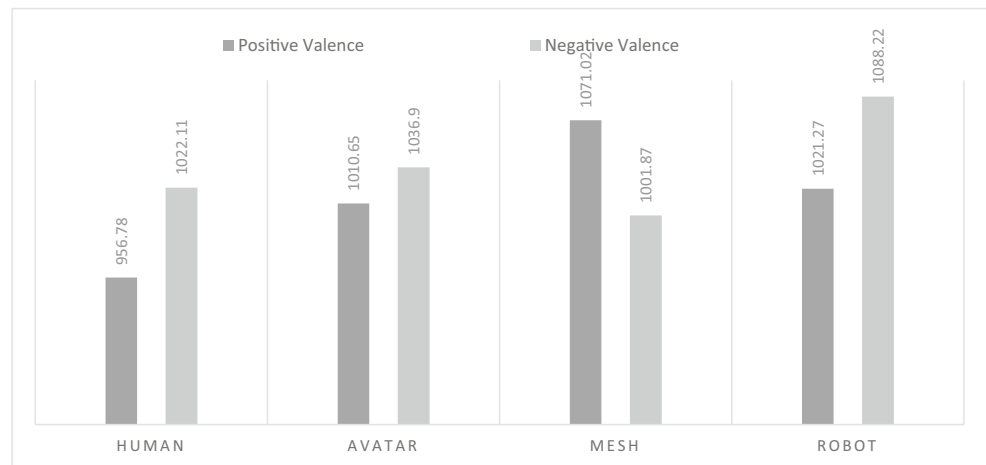
t1.1 **Table 1** Comparison of means  
t1.2 according to the type of face for  
response time

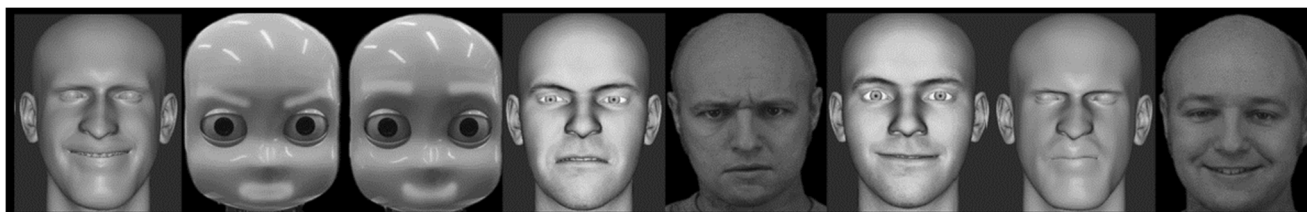
	Means	SD	Statistic (Z)	p value	Effect size	
t1.3	Sample					
t1.4	Human/avatar	1978.89/2047.55	369.70/446.08	-2.048	<b>0.041*</b>	0.167
t1.5	Human/mesh	197,889/2072.89	369.70/453.169	-2.888	<b>0.004*</b>	0.227
t1.6	Human/robot	1978.89/2109.49	369.70/472.022	-3.814	<b>0.000*</b>	0.308
t1.7	Avatar/mesh	2047.55/2072.89	446.08/453.169	-1.171	0.241	0.057
t1.8	Avatar/robot	2047.55/2109.49	446.08/472.022	-2.361	<b>0.018*</b>	0.135
t1.9	Mesh/robot	2072.89/2109.49	453.169/472.022	-0.445	0.656	0.079
t1.10	Positive valence condition					
t1.11	Human/avatar	956.78/1010.65	184.11/239.28	-1.966	<b>0.049*</b>	0.252
t1.12	Human/mesh	956.78/1071.02	184.11/301.68	-4.032	<b>0.000*</b>	0.457
t1.13	Human/robot	956.78/1021.27	184.11/263.72	-2.029	<b>0.042*</b>	0.283
t1.14	Avatar/mesh	1010.65/1071.02	239.28/301.68	-2.888	<b>0.004*</b>	0.221
t1.15	Avatar/robot	1010.65/1021.27	239.28/263.72	-0.867	0.386	0.042
t1.16	Mesh/robot	1071.02/1021.27	301.68/263.72	-1.898	0.058	0.175

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

compared with responses given for human faces. More partic-  
ularly, the discrimination of positive expressions requires sig-  
nificantly 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, imag-  
ine 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 expres-  
sions than to positive valence expressions (Pais et al. 2013).  
These results are consistent with a recent meta-analysis con-  
ducted 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)

309 happy faces are better recognized in photographs, whereas  
 310 angry/sad faces are recognized better in schematic faces.  
 311 This notion seems to be the explanatory key of the results  
 312 obtained for the “mesh” face. Indeed, negative emotions were  
 313 identified faster than positive emotions. Some studies have  
 314 shown that when a face is ambiguous it is attributed to a more  
 315 negative valence than when it is less ambiguous (Cheetham  
 316 et al. 2015; Lazzeri et al. 2015). This may be why we obtained  
 317 a lesser number of correct responses for this face type.  
 318 Participants interpret it more negatively and attribute a nega-  
 319 tive valence even for expressions with a positive valence.  
 320 Authors explain this tendency by our innate predisposition  
 321 to treat strangers with caution (Zajonc 1998). Moreover, the  
 322 skin reflects our identity, health care, and other individual  
 323 characteristics (Alkawaz and Basori 2012). This is what  
 324 makes us human, and warm. For example, the lack of wrinkles  
 325 affects emotion perception and recognition (Milcent et al.  
 326 2019). We can assume that the mesh stimulus does not allow  
 327 the collection of relevant information. It would then be con-  
 328 sidered as an ambiguous stimulus treated negatively by the  
 329 participants. The realism needed to ward off the ambiguity  
 330 of facial expressions may lead to the uncanny valley issue.  
 331 There comes a time when if the agent looks too similar while  
 332 retaining certain flaws, these appear to be particularly mon-  
 333 strous and disturbing, as when observing a zombie, for exam-  
 334 ple. The degree of familiarity contrasts too starkly with the  
 335 flaws that remind us that the agent is not a living being when  
 336 human behavior is expected from him. However, beyond this  
 337 valley, familiarity reaches a second peak, and humanoids are  
 338 accepted again. This theory, although controversial, has had a  
 339 huge impact on robotics and 3D animation. Designers are  
 340 advised not to attempt to reach the second familiarity peak  
 341 present after the uncanny valley given the risk of not reaching  
 342 it (MacDorman 2019) It is because they do not claim to reach  
 343 the second peak that the avatars presented in our study were  
 344 able to achieve such good results. Identifying them as a hu-  
 345 manoid agent was enough to understand them and interpret  
 346 their facial expressions. This maintained a relevant degree of  
 347 affinity and familiarity. The use of this type of face seems  
 348 relevant for social skills training or every training with non-  
 349 human agents. Individuals recognize and understand these  
 350 agents as well as human faces. Moreover, clinical populations  
 351 with social cognition deficit show altered recognition skill of

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

### 367 Limitations

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

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



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

395 **Conclusion**

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

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

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

417 **Compliance with Ethical Standards**

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

420 **References**

421 Adolphs, R. (1999). Social cognition and the human brain. *Trends in*  
 422 *Cognitive Sciences*, 3(12), 11.  
 423 Adolphs, R., Baron-Cohen, S., & Tranel, D. (2002). Impaired recognition  
 424 of social emotions following amygdala damage. *Journal of*  
 425 *Cognitive Neuroscience*, 14(8), 1264–1274. <https://doi.org/10.1162/089892902760807258>.  
 426 Alkawaz, M. H., & Basori, A. H. (2012). The effect of emotional colour  
 427 on creating realistic expression of avatar. Proceedings of the 11th  
 428 ACM SIGGRAPH international conference on virtual-reality contin-  
 429 uum and its applications in industry - VRCAI '12, 143. <https://doi.org/10.1145/2407516.2407555>.  
 430 Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003).  
 431 Interpersonal distance in immersive virtual environments.  
 432 *Personality and Social Psychology Bulletin*, 29(7), 819–833.  
 433 <https://doi.org/10.1177/0146167203029007002>.  
 434 Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison  
 435 of Beck depression inventories-IA and-II in psychiatric outpatients.  
 436 *Journal of Personality Assessment*, 67(3), 588–597. [https://doi.org/10.1207/s15327752jpa6703\\_13](https://doi.org/10.1207/s15327752jpa6703_13).  
 437 Biocca, F., Harms, C., & Burgoon, J. K. (2003). Toward a more robust  
 438 theory and measure of social presence: Review and suggested  
 439

criteria. *In Presence: Teleoperators and Virtual Environments*, 442  
 443 456–480.  
 Breazeal, C. (2003). Emotion and sociable humanoid robots. 444  
*International Journal of Human-Computer Studies*, 59(1–2), 119– 445  
 155. [https://doi.org/10.1016/S1071-5819\(03\)00018-1](https://doi.org/10.1016/S1071-5819(03)00018-1). 446  
 Cheetham, M., Wu, L., Pauli, P., & Jancke, L. (2015). Arousal, valence, 447  
 and the uncanny valley: Psychophysiological and self-report find- 448  
 ings. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00981>. 449  
 Churamani, N., Cruz, F., Griffiths, S., & Barros, P. (2016). iCub: 450  
 Learning emotion expressions using human reward. 451  
 Dubal, S., Foucher, A., Jouvent, R., & Nadel, J. (2011). Human brain spots 452  
 emotion in non humanoid robots. *Social Cognitive and Affective* 453  
*Neuroscience*, 6(1), 90–97. <https://doi.org/10.1093/scan/nsq019>. 454  
 Dyck, M., Winbeck, M., Leiberg, S., Chen, Y., Gur, R. C., & Mathiak, K. 455  
 (2008). Recognition profile of emotions in natural and virtual faces. 456  
*PLoS One*, 3(11), e3628. <https://doi.org/10.1371/journal.pone.0003628>. 457  
 Dyck, M., Winbeck, M., Leiberg, S., Chen, Y., & Mathiak, K. (2010). 458  
 Virtual faces as a tool to study emotion recognition deficits in 459  
 schizophrenia. *Psychiatry Research*, 179(3), 247–252. <https://doi.org/10.1016/j.psychres.2009.11.004>. 460  
 Ekman, P. (1993). Facial expression and emotion. *American* 461  
*Psychologist*, 48(4), 384–392. <https://doi.org/10.1037/0003-066X.48.4.384>. 462  
 Eyssel, F., Hegel, F., Horstmann, G., & Wagner, C. (2010). 463  
 Anthropomorphic inferences from emotional nonverbal cues: A 464  
 case study. *19th International Symposium in Robot and Human* 465  
*Interactive Communication*, 646–651. <https://doi.org/10.1109/ROMAN.2010.5598687>. 466  
 Frith, C. (2009). Role of facial expressions in social interactions. 467  
*Philosophical Transactions of the Royal Society B: Biological* 468  
*Sciences*, 364(1535), 3453–3458. <https://doi.org/10.1098/rstb.2009.0142>. 469  
 Hall, J., Tritton, T., Rowe, A., Pipe, A., Melhuish, C., & Leonards, U. 470  
 (2014). Perception of own and robot engagement in human–robot 471  
 interactions and their dependence on robotics knowledge. *Robotics* 472  
*and Autonomous Systems*, 62(3), 392–399. <https://doi.org/10.1016/j.robot.2013.09.012>. 473  
 Hortensius, R., Hekele, F., & Cross, E. S. (2018). The perception of 474  
 emotion in artificial agents. *IEEE Transactions on Cognitive and* 475  
*Developmental Systems*, 10(4), 852–864. <https://doi.org/10.1109/TCDS.2018.2826921>. 476  
 Jack, R. E., & Schyns, P. G. (2017). Toward a social psychophysics of 477  
 face communication. *Annual Review of Psychology*, 68(1), 269– 478  
 297. <https://doi.org/10.1146/annurev-psych-010416-044242>. 479  
 Joyal, C. C., Jacob, L., Cigna, M.-H., Guay, J.-P., & Renaud, P. (2014). 480  
 Virtual faces expressing emotions: An initial concomitant and con- 481  
 struct validity study. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00787>. 482  
 Kahn Jr., P. H., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., 483  
 Gill, B. T., Ruckert, J. H., & Shen, S. (2012). “Robovie, you’ll have 484  
 to go into the closet now”: Children’s social and moral relationships 485  
 with a humanoid robot. *Developmental Psychology*, 48(2), 303– 486  
 314. <https://doi.org/10.1037/a0027033>. 487  
 Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and pu- 488  
 berty influence the development of facial emotion recognition. 489  
*Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00761>. 490  
 Lazzari, N., Mazzei, D., Greco, A., Rotesi, A., Lanatà, A., & De Rossi, D. 491  
 E. (2015). Can a humanoid face be expressive? A psychophysiological 492  
 investigation. *Frontiers in Bioengineering and Biotechnology*, 3. 493  
<https://doi.org/10.3389/fbioe.2015.00064>. 494  
 Leppänen, J. M., & Hietanen, J. K. (2004). Positive facial expressions are 495  
 recognized faster than negative facial expressions, but why? 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506

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

## 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.