



**HAL**  
open science

# Combining trunk movement and facial expression enhances the perceived intensity and believability of an avatar's pain expression

Thomas Treal, Philip L. Jackson, Aurore Meugnot

► **To cite this version:**

Thomas Treal, Philip L. Jackson, Aurore Meugnot. Combining trunk movement and facial expression enhances the perceived intensity and believability of an avatar's pain expression. *Computers in Human Behavior*, 2020, 112, pp.106451 -. 10.1016/j.chb.2020.106451 . hal-03492365

**HAL Id: hal-03492365**

**<https://hal.science/hal-03492365>**

Submitted on 18 Jul 2022

**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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

## **Combining trunk movement and facial expression enhances the perceived intensity and believability of an avatar's pain expression**

Thomas Treal<sup>a,b</sup>, Philip L. Jackson<sup>c,d,e</sup>, Aurore Meugnot<sup>a,b,\*</sup>

<sup>a</sup> Université Paris-Saclay CIAMS, 91405, Orsay, France

<sup>b</sup> CIAMS, Université d'Orléans, 45067, Orléans, France

<sup>c</sup> École de Psychologie, Université Laval, Québec, Canada

<sup>d</sup> Centre Interdisciplinaire de Recherche en Réadaptation et Intégration Sociale (CIRRIS), Québec, Canada

<sup>e</sup> CERVO Research Center, Québec, Canada

\*Corresponding author: Faculté des sciences du sport de l'Université Paris Sud, Bât. 335 rue Pierre de Coubertin, 91405 Orsay France

*E-mail address* : [aurore.meugnot@universite-paris-saclay.fr](mailto:aurore.meugnot@universite-paris-saclay.fr)

# Combining trunk movement and facial expression enhances the perceived intensity and believability of an avatar's pain expression

## 1. Introduction

The last decade has witnessed a great enthusiasm for the use of computer-generated (CG) characters (avatars) in social and affective neuroscience studies. This type of technology promises to meet the idealistic compromise between controlled and ecological paradigms (Bombari, Schmid Mast, Canadas & Bachman, 2015; Pan & Hamilton, 2018; Wykowska, Chaminade, Cheng, 2016). That being said, the development of platforms allowing the animation of avatars remains complex and has given rise to the key issue of their *realism* (i.e., the degree of human likeness of the virtual character), which may be crystallized through the 'fear' of falling into the *Uncanny Valley* phenomenon (Kätsyri, de Gelder & Takala, 2019; MacDorman, Green, Ho & Koch, 2009; Mathur & Reichling, 2016; Reuten, Van Dam & Naber, 2018; see Brenton, Gillies, Ballin & Chatting, 2005; Kätsyri, Förger, Mäkäräinen, & Takala, 2015 for critical reviews). This well-known phenomenon refers to 'a profound sense of unease provoked by an almost but not fully realistic character that prevents any sense of presence or emotional engagement' (Mori, 1970/2012). In the domain of emotion research, the use of anthropomorphic avatars may be crucial. If the avatar is perceived as a simple machine, its facial emotional expression would not be *believable*. It would be considered as fake and artificial and would unlikely trigger any spontaneous emotional reaction (Bogdanovych, Trescak & Simoff, 2016). Such an experimental paradigm would not be useful in the context of inferring the mechanisms underlying others' emotion processing and regulation.

Scholars argued that the lack of believability towards the avatar's behavior, and the eerie feeling that may accompany it, would be elicited by a perceptual mismatch in the realism of the avatar's features, with some appearing human and others very artificial (e.g., artificial eyes on a realistic face or disparity between voice and face realism levels, MacDorman et al., 2009; Kätsyri et al., 2015). In particular, realism inconsistency could result from a gap between a highly graphically realistic avatar and a low realism in terms of motion and behavior (Brenton et al., 2005; de Gelder, Kätsyri & de Borst, 2018). This discrepancy could be the result of the disproportional amount of efforts put into pictural (or graphical) realism in avatar development (Brenton et al., 2005) as opposed to their movements -- perhaps due to the difficulty of gesture animation (Hartmann, Mancini & Pelachaud, 2005; Normoyle, Lui, Kapada, Badler & Jorg, 2013). This is nevertheless quite surprising considering the bulk of evidence supporting the role of body expression in conveying emotion (de Gelder, 2006; de Gelder, de Borst & Watson, 2015). It is even more surprising when one considers that Mori - who is extensively cited in the field of affective computing - attributed the highest level of affinity with an avatar that portrays a moving healthy character (Mori, 1970; 2012). One might expect this observation to instill momentum into the development of both the pictural and behavioral features of virtual characters, notably when using them as stimuli to explore the interactive mechanisms underlying emotion perception and regulation.

To our knowledge, studies exploring the influence of body motion on the perception of facial emotions 'expressed' by an avatar are scarce (Buisine et al., 2014; Clavel, Plessier, Martin & Morel, 2009; De Melo, Kenny & Gratch, 2010; Wu et al., 2014). For instance, in an elegant study, De Melo and al. (2010) demonstrated that implementing respiratory patterns corresponding to various emotions (excitement, pain, relief, anger, fear, panic, boredom and startle) to an avatar increased the perception of these emotions in a virtual character, compared with a control (avatar without body animation) condition. In another seminal work,

Clavel and collaborators (2009) examined how body posture may contribute to the recognition of basic emotion (fear, anger, joy, surprise, sadness) expressed by a virtual character. The authors showed that participants recognized these emotions more easily when the avatar's face and posture were jointly presented, compared to conditions with either face or posture alone (Clavel et al., 2009). The same group replicated similar results in scenarios of an avatar expressing emotions while telling a story including ironic and narrative items (Buisine et al. 2014). The authors showed that the negative emotions were judged to be more intense when the avatar adopted a posture in congruence with its facial expression compared to a condition where the avatar was still. Interestingly, Buisine and collaborators (2014) also examined the influence of an *idle* posture (i.e. small variations in body postures or balance changing) on participants' perception of the avatar's emotions, compared with a fixed posture. *Idle motion* was coined by Egges (2004) to describe neutral (i.e. non-communicative) movements, including "*changing balance because of fatigue, small variation in body posture caused by small muscle contraction, eye blinking*". In other terms, idle motion refers to subtle movements that are not always perceptible if one pays no attention to them. Such movements are almost always present in humans (Perlin, 1995). These movements are generally lacking in avatar animation, but they might be a core ingredient to creating the illusion that the avatar is alive. There are some exceptions, for instance idle motion implemented by default in a conversational avatar in Groom et al. (2009), and the control condition in De Melo and al., (2010). Buisine and al. (2014) observed that participants tend to recognize more emotions in the idle than in the still condition, while there was no significant difference between the two conditions for the perceived emotion intensity, as well as the evaluation of emotion realism. Although mixed, these results are important to consider because, to our knowledge, this is the first study examining how this type of body motion - which may be a critical component of the avatar's animation - may influence the perception of emotion as well as its overall realism.

Moreover, other authors showed that participants' emotional reactions were stronger when viewing the pain of a virtual patient whose medical condition was growing worse in a dynamic condition (the avatar is moving) compared to a condition where the virtual patient remained static (Wu et al., 2014). This latter study may be of particular interest for the present work, which has used a similar virtual reality platform using recently avatars expressing pain in a medical context (Tessier, Gingras, Robitaille & Jackson, 2019). Finally, authors investigated the psychophysiological response to emotional stimuli by the means of pupil size changes (Bradley, Miccoli, Escrig & Lang, 2008; Kret, Roelofs, Stekelenburg & de Gelder, 2013; Reuten, Van Dam & Naber, 2018). Bradley and al. (2008) observed that pupil diameter was larger when participants were presented with pictures eliciting both pleasant and unpleasant contents than pictures with a neutral content. According to the authors, pupil dilation may be caused by sympathetic nervous system activity and reflect emotional arousal, which is a dimension of emotional reaction. Indeed, emotional reaction is often defined as the combination of two systems: the emotional valence (pleasantness/positive vs unpleasantness/negative), which indicates the direction of the emotional reaction, and the emotional arousal, which refers to the intensity of the emotional reaction (Lang, Bradley, & Cuthbert, 1990; Russell, 1979). Thus, pupil dilation is well established evidenced as a continuous and objective marker of emotional arousal when viewing facial expressions. Other authors recently showed that pupillary responses may also change when observing a virtual character's facial expression . Pupil responded less vigorously if participants judged the avatar's facial expression as eerie or unattractive (Reuten, van Dam & Naber, 2018).

Thus, two studies were conducted. Experiment 1 aimed to determine the impact of different levels of trunk movement on *a) the perceived intensity* of pain expressed by an avatar, *b) the believability* of its facial pain expression, and *c) the participant's emotional arousal* assessed by pupil size recording. It was hypothesized that the implementation of trunk

movement with a low amplitude (i.e., idle motion representing postural oscillation, typical of human dynamic equilibrium) would improve participants' perception of the avatar's pain and believability compared to a static condition. Larger amplitudes of trunk oscillation (i.e. trunk rocking) were supposed to express a pain behavior and would strengthen the congruency between face and body. Consequently, the avatar's pain expression would be intensified, up to a certain point where trunk oscillation would seem exaggerated and might make the avatar's pain seem weird, so less believable. Moreover, the pupil would dilate more if the avatar's pain were perceived to be more believable and more intense, namely in the conditions where trunk movement has been added to it compared with the static condition.

Experiment 2 using psychophysical cues was conducted to objectively classify the different conditions of trunk movement as idle motion or trunk rocking. The results observed in the Experiment 1 regarding the impact of trunk movement on participants' perception of the avatar's pain (*Pain Intensity* and *Pain Believability*) and the emotional arousal elicited by pain (*Pupil Size Change*) will be further discussed in light of the categorization obtained in Experiment 2.

## **2. Experiment 1: Effect of body movement added to an avatar on the way one perceives facial expression of pain.**

### ***2.1. Method***

#### ***2.1.1. Participants***

Twenty-four participants (mean age = 22.04 years, SD = 3.61, range = 18-33 years; 13 men (54.2%)) were recruited through advertisements sent via South-Paris University's student mailing lists or posted on the campus of South-Paris University. Exclusion criteria included any reported history of painful, neurological, psychiatric disorder, work experience as a healthcare provider and an uncorrected visual impairment. This study was approved by the

Ethics Committee of Paris-Saclay University (#CER-Paris-Saclay-2018-16) and participants gave written informed consent to take part in this study.

### **2.1.2. Procedure**

Participants sat comfortably on a chair with their head supported by a chin-rest and facing a 20-inch monitor (60Hz refresh rate, resolution 1680 x 1050 pixels) at a viewing distance of 55cm. A 9-point calibration of the eye-tracker was conducted until a good calibration was accepted (i.e. with a maximum error angle of 1°). After the calibration phase, participants were instructed to not move their head until the end of the experiment. Then, they completed *the Pain Intensity task* and *the Pain Believability task*. The order of the tasks was counterbalanced between the participants. After the experiment, a debriefing was held to inform the participants of the true objective of the study.

### **2.1.3. Material and tasks**

Seven video clips (duration = 4 seconds) displaying the upper body from the trunk up of a male avatar originated from the Empathy-Enhancing Virtual Evolving Environment (EEVEE) (Jackson et al., 2015) were created with Blender® (BlenderFoundation, Amsterdam, Netherlands). The avatar facing the camera at an 86.8° angle showed a dynamic expression of pain, either with or without an oscillatory movement of the trunk.

The avatar's pain expression was created using the Action Units (AUs) according to the Facial Action Coding System (FACS) (Ekman & Friesen 1978), which are typical of pain expression (Prkachin, 1992): AU4 (brow lower), AU6 (cheek raiser), AU7 (lid tightener), AU9 (nose wrinkle), AU10 (upper lip raiser) and AU43 (eyes closed); and five occasionally

AUs: AU12 (lip corner puller), AU20 (lip stretch), AU25 (lips part), AU26 (jaw drop) and AU27 (mouth stretch). The intensity of the pain expression was fixed at 60% of the maximal contraction of the AUs described above. Each video began with the avatar showing a neutral expression (sets of AUs at 0). The level of AUs linearly increased to reach 60% of contraction for 1 second. The pain expression was maintained for 3 seconds until the end of the clip (Figure 1).



**Figure 1 – Left figure: Avatar at the onset of the clip (neutral expression). Right figure: Avatar expressing pain (AUs recruited at 60% of their maximal contraction)**

In six video clips, the avatar's trunk was moving to and fro at different amplitudes ( $0.33^\circ$ ;  $0.67^\circ$ ;  $1^\circ$ ;  $1.67^\circ$ ;  $2^\circ$ ;  $3^\circ$  conditions). Trunk movement was limited to the anterior-posterior part and the degree of oscillation was determined to be at some point between the center of the mass and the top of the head. In each video, the avatar first stood perpendicular to the ground and moved forward (half amplitude), then moved back (total amplitude) and finally returned to its initial position (half amplitude) (Figure 2). In the seventh clip, no trunk movements were added to the avatar (*still condition*).



**Figure 2: illustration of the trunk movement implemented on the avatar**

Thus, the stimuli were composed of seven video clips of a male avatar expressing pain at 60 % of pain-related AUs maximal contraction, with different conditions of trunk movement ( $0^\circ$ ,  $0.33^\circ$ ;  $0.67^\circ$ ;  $1^\circ$ ;  $1.67^\circ$ ;  $2^\circ$ ;  $3^\circ$ ). The video clips are available online with a link placed in the section Supplementary material. Two tasks were created with these video clips and they were displayed using Experiment Builder (SR Research, Canada):

*The Pain Intensity task.* Participants had to evaluate the intensity of the avatar's pain expression by moving a cursor on a visual analog scale (VAS) (Left extremity = "No Pain", Right extremity = "Worst Pain Imaginable", [in French: "Pas de douleur" and "Pire douleur imaginable"]). Each clip was displayed 10 times during the experimental phase. In total, this phase was composed of 70 randomized trials. A trial began with a centered fixation cross which was displayed for 1500ms. Then, the video clip was presented (duration=4 seconds), followed by the VAS. Every 10 trials, a break-screen was displayed until participants pressed the space key to continue the task. Before the experimental phase, the participants performed a familiarization phase composed of three trials with a break-screen between the second and third familiarization trials.

*The Pain Believability task.* Participants had to judge whether the avatar displayed in the clips was genuinely experiencing pain or not. To this end, they were informed that the avatar was animated with a pain expression captured either on individuals who were actually experiencing pain ("True pain") or actors imitating a pain behavior ("Fake pain"). Like in the

pain intensity task, this task was composed of a familiarization phase with three stimuli and a break-screen, followed by an experimental phase of the seven video clips presented 10 times at random (70 trials in total). A trial began with a centered cross fixation lasting 1500 ms. Then, a video clip was presented followed by a screen displaying two panels: “True pain” and “Fake pain”, respectively centered on the left and the right part of the screen [in French: "VRAI" and "FAUX"]. To respond, participants had to click on one panel with a mouse.

#### ***2.1.4. Eye-tracking***

Pupil size was collected continually during the experiment using a monocular desk-mounted eye-tracker (EyeLink 1000, SR Research, Canada) at a rate of 1000 Hz and registered in arbitrary units using Experiment Builder (SR Research, Canada).

#### ***2.1.5. Data Analysis***

*The Pain Intensity task.* The position of the cursor along the scale was converted to numerical values between 0 (“no pain”) and 100 (“worst pain imaginable”). The mean *Pain Intensity* was calculated for each condition of trunk movement (condition 0° included).

*The Pain Believability task.* The *Pain Believability* rating corresponding to the percentage of “True pain” responses was calculated for each condition of trunk movement (condition 0° included).

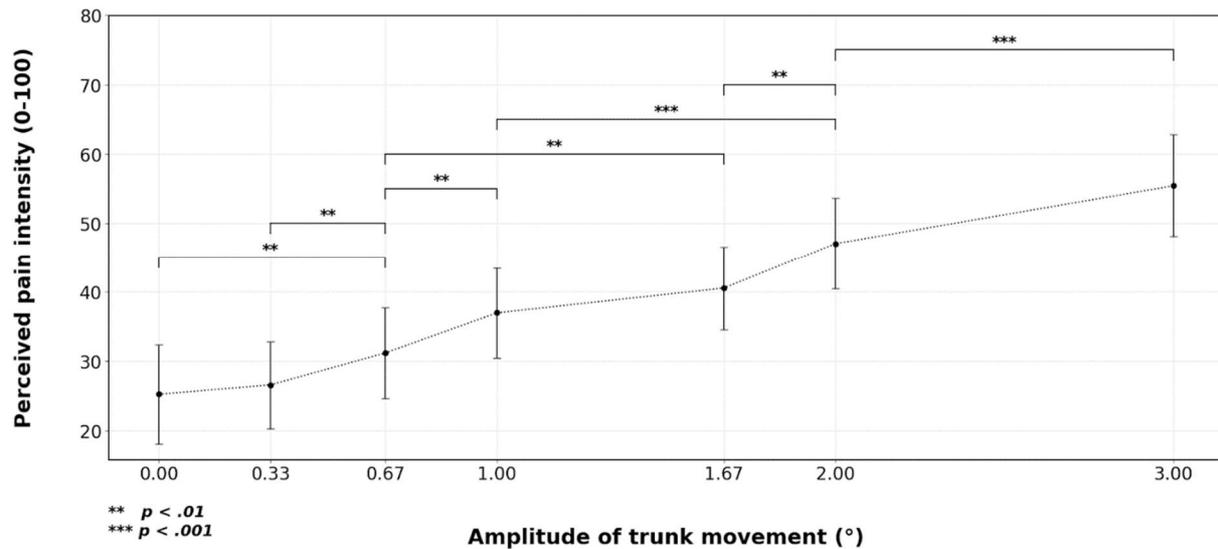
*Eye-tracking.* *Pupil size* was analyzed using DataViewer (SR Research, Canada). One participant was removed from the analysis because of a high data loss during eye tracking recording (37.5%, mean data loss for others: 1.7% (SD=1.8). The data were epoched relative to the onset and end of the video clips (70 epochs x 2 tasks = 140 epochs of 4 seconds for each participant). The first second of each epoch was removed from the analysis to exclude the change related to the initial pupillary light response (Mitz, Chacko, Putnam, Rudebeck, & Murray, 2017). The average pupil size for each condition of trunk oscillation was then

calculated between the 1- and 4-second marks. A change rate (%) in pupil size was calculated for the six conditions of trunk movement relative to the static condition ( $0^\circ$ ) using the following formula: (mean pupil size for given condition \* 100 / mean pupil size for condition  $0^\circ$ ) – 100. A positive change indicated a larger pupil size in the condition of trunk movement than in the condition  $0^\circ$ . A negative change indicated a smaller pupil size in the trunk movement than in the condition  $0^\circ$ .

All statistical analyses were conducted using SPSS. Analysis of variance (ANOVAs) with repeated-measures were performed on both *Pain Intensity* (0-100) and *Pain Believability* (%) with the condition of trunk oscillation ( $0^\circ$ ,  $0.33^\circ$ ;  $0.667^\circ$ ;  $1^\circ$ ;  $1.667^\circ$ ;  $2^\circ$ ;  $3^\circ$ ) as within-subject factor. A Greenhouse-Geisser correction was used when the sphericity assumption was violated ( $p < .05$  for the Mauchly's test of sphericity). Alpha was set at .05 and in case of significance, pairwise comparisons using Bonferroni correction were applied. T-tests relative to 0 were performed on *Pupil Size Change* for each condition of trunk oscillation ( $0.33^\circ$ ;  $0.67^\circ$ ;  $1^\circ$ ;  $1.67^\circ$ ;  $2^\circ$ ;  $3^\circ$ ), with alpha set at 0.5. Pairwise comparisons were applied between conditions with a significant difference from 0 using Bonferroni correction if required.

## **2.2. Results**

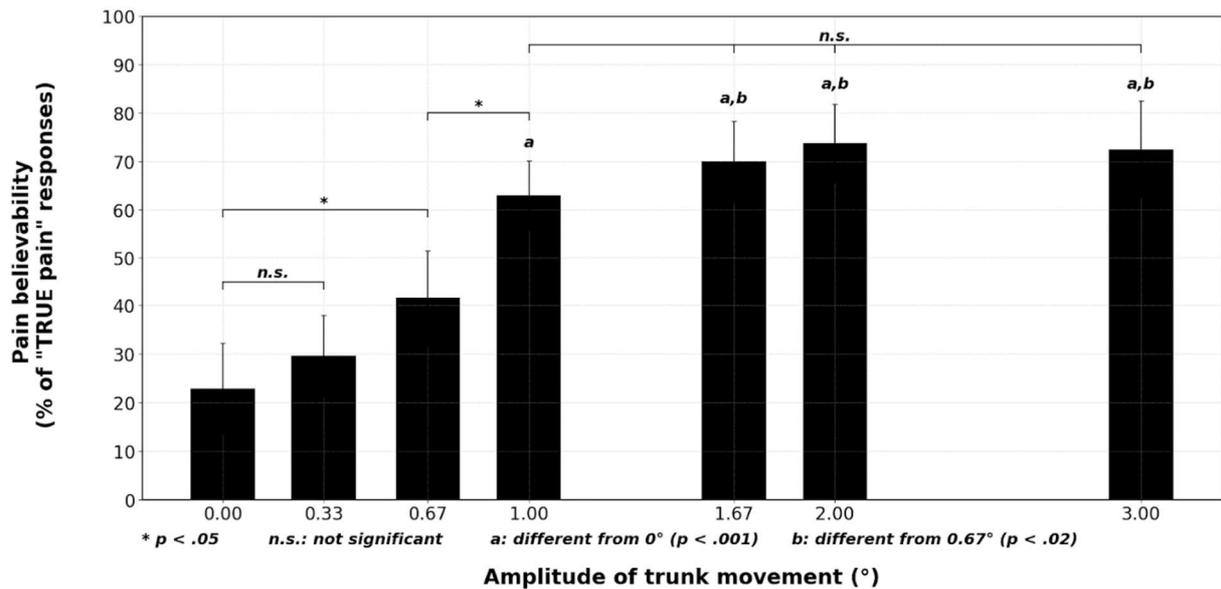
*The Pain Intensity task.* The repeated-measures ANOVA revealed a large and significant effect of trunk movement on the evaluation of *Pain Intensity* [ $F(1.5,35) = 43.54$ ,  $p < .001$ ,  $\eta^2 = .30$ ,  $\eta p^2 = .65$ ]. As illustrated in Figure 3, pairwise comparisons with Bonferroni correction indicated that the judgment of *Pain Intensity* increased significantly with the amplitude of trunk movement, except between the lower amplitude of oscillation ( $0.33^\circ$ ) compared to the condition without oscillation ( $0^\circ$ ) ( $p = 1.0$ ).



**Figure 3 – Perceived Pain Intensity of facial pain expression according to the amplitude of trunk movement.**

Error bars correspond to the confidence intervals at 95% (using t table and corrected standard deviation because  $N < 30$ ). All significance thresholds were adjusted using Bonferroni correction.

*The Pain Believability task.* The repeated-measures ANOVA revealed a large and significant effect of trunk movement on the evaluation of *Pain Believability* [ $F(2.737, 62.948) = 23.16, p < .001, \eta^2 = .47, \eta p^2 = .50$ ]. As illustrated in Figure 4, pairwise comparisons with Bonferroni correction indicated that the avatar's pain expression was significantly judged to be more believable than the condition ( $0^\circ$ ) (i.e. higher % of true) across all conditions [ $0.67^\circ (p = .022, d = .76), 1^\circ (p < .001, d = 1.43), 1.67^\circ (p < .001, d = 1.16), 2^\circ (p < .001, d = 1.38)$  and  $3^\circ (p < .001, d = 1.21)$ ] except lower condition of oscillation ( $0.33^\circ; p = 1.0$ ). The condition  $1^\circ$  was judged to be significantly more believable than the condition  $0.67^\circ (p = .04, d = .72)$ . Moreover, there was no significant difference between conditions  $1^\circ, 1.67^\circ, 2^\circ$  and  $3^\circ (p > .378)$ , indicating the reach of a ceiling.



**Figure 4 – Pain Believability rating according to the amplitude of the oscillations**

Error bars correspond to the confidence intervals at 95% (using t table and corrected standard deviation because  $N < 30$ ). All significance thresholds were adjusted using Bonferroni correction.

*Eye-tracking.* As illustrated in Figure 5 and as was confirmed by the t-test relative to 0, the *pupil size* was significantly larger from the static condition ( $0^\circ$ ) for the conditions with an amplitude of  $1^\circ$  ( $p = .014$ ,  $d = .55$ ) and  $3^\circ$  ( $p = .012$ ,  $d = .57$ ) of oscillation. Paired t-test between the conditions  $1^\circ$  and  $3^\circ$  did not reveal any significant differences [ $t(22) = -1.414$ ,  $p = .171$ ,  $d = -.295$ ] suggesting the reach of a ceiling for pupil size after  $1^\circ$  of oscillation.

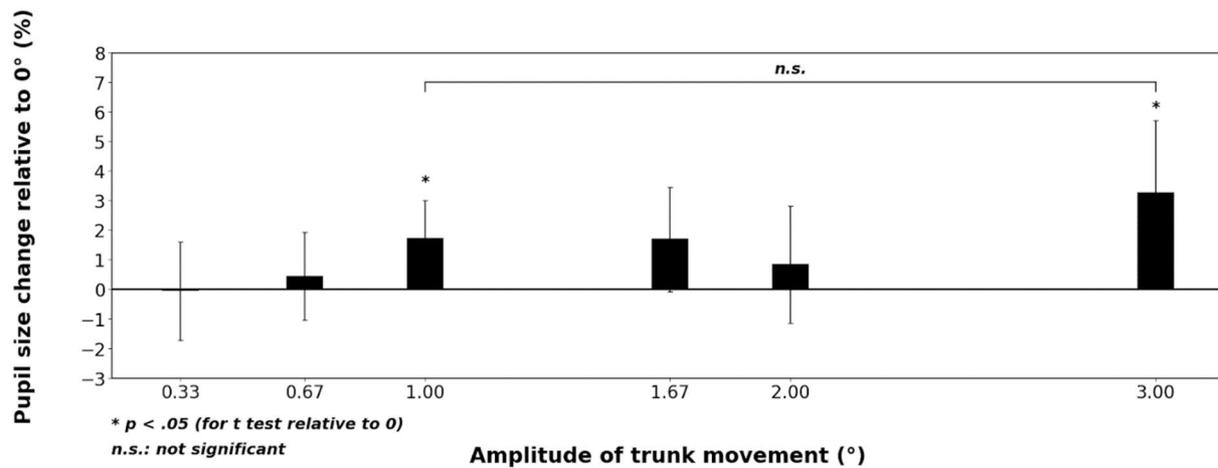


Figure 5 – Mean of pupil size change, relative to 0° (%) according to the amplitude of trunk movement. Error bars correspond to the confidence intervals at 95% (using  $t$  table and corrected standard deviation because  $N < 30$ )

### 2.3. Summary of Experiment 1

Participants rated the avatar’s pain expression as gradually more intense (condition 0°: mean score = 25.21, 95% CI [18.06, 32.36]; condition 3°: mean score 55.38, 95% CI [48.02, 62.74]). Note that the average score of pain intensity in the condition 3° was similar to the level of pain expression that was modeled based on levels of pain-related AUs (i.e., 60 % of their maximal contraction). The avatar’s pain expression was also perceived as more believable when the avatar was animated with trunk movement (0°: mean rating = 22.92%, 95% CI [13.55, 32.29]; 1°- 3°: mean rating = 62.92%, 70%, 73.75%, 72.5%; 95% CIs [55.8, 70.04], [61.72, 78.28], [65.66, 81.84], and [62.48, 82.52] respectively). Contrary to our hypothesis, the largest amplitude of trunk oscillation (3°) did not provoke a decrease in the perceived intensity and believability of pain in the avatar compared with the still and lower conditions of trunk movement. Unexpected results were also observed for pupillary response . The pupil dilated more strongly only when the avatar oscillated at a large amplitude (conditions 1°: 1.71%, 95% CI [0.41, 3.01] and 3°: 3.28%, 95% CI [0.85, 5.71]) than when it

stayed still. Surprisingly, there was a lack of significant change for the conditions  $1.67^\circ$  and  $2^\circ$  compared with the static condition, while there was a significant change for the preceding ( $1^\circ$ ) and next ( $3^\circ$ ) conditions (Figure 5). Importantly, the lower amplitude of trunk movement (condition  $0.33^\circ$ ) did not differ from the still condition for the three dependent variables (i.e. *Pain Intensity*, *Pain Believability* and *pupil size*). This lack of difference may suggest that this amplitude was too low to be perceived, and raised the question of its definition as idle motion.

A second study (*Experiment 2*) was conducted to objectively classify the different conditions of trunk movement as idle motion or trunk rocking, which may reflect a pain behavior when combined with facial pain expression. The results observed in *Experiment 1* regarding the impact of trunk movement on participants' perception of the avatar's pain and the emotional arousal elicited by pain will be further discussed in light of the categorization obtained in *Experiment 2*.

### **3. Experiment 2: Categorization of trunk movement as idle motion or trunk rocking**

A classical psychophysical tool was used to determine the categorical boundaries between the different conditions of trunk movement (Bonnet, 1986; Courbalay et al., 2015; Wichmann & Hill, 2001). The aim was to classify them as idle motion or trunk rocking, which may reflect a pain behavior in the context of this study (i.e. when combined with facial pain expression). The video clips used in *Experiment 1* were presented to a group of naive participants in a visual discrimination task. [Herein](#), the participants were asked whether or not they had perceived the avatar's chest moving in each video clip. To be considered an idle motion, a trunk movement must be perceived (i.e. above a *detection threshold*), yet not systematically (i.e., under a *systematic detection threshold*), as these movements are by definition subtle and thus difficult to detect. The boundaries of idle motion were then calculated by fitting a psychophysical function to the detection rating of each amplitude of

trunk movement (see below *section 3.1.2* for an explanation of the related psychophysical indices). The amplitudes of trunk oscillation which fell above the *systematic detection threshold* were considered as the bodily expression of a sustained pain.

### **3.1. Method**

#### **3.1.1. Participants**

The participants were twenty young adults recruited in the Faculty of Sport Science of South-Paris University (11 women, range = 18-27 years, mean age =  $20.3 \pm 2.65$  years) through advertisements sent via South-Paris University's student mailing lists or a course program which proposed, on a voluntary basis, to obtain one extra mark on a test when participating to an experiment conducted in the laboratory. All participants had (or were corrected to) normal vision. This study was approved by the Ethics Committee of Paris-Saclay University (#CER-Paris-Saclay-2018-16) and participants gave written informed consent to take part in this study.

#### **3.1.2. Material and Task**

The same video clips as in the first experiment were presented to participants sat in front of the computer screen (20-inch, 60Hz refresh state, resolution = 1680 x 1050 pixels) at a distance of 55 cm from the head. The clips were presented 10 times per condition (70 randomized trials). A trial began with the display of a clip, and after participants had to respond whether the avatar's trunk moved or not by clicking "YES" or "NO" [in French, "OUI" or "NON"] labels, respectively presented on the center-left and center-right of the screen .

### 3.1.3. Data Analysis

A logistic function  $y = 1/(1+\exp(-(x-PSE)/JND))$  was fitted to the average of the percentage of “YES” responses for each condition. The Point of Subjective Equality (PSE) corresponds to the detection threshold (the lowest amplitude of trunk movement that participants are able to detect). The Just Noticeable Difference (JND) refers to the precision of categorization. It corresponds to the minimum amplitude necessary to perceptually differentiate two sizes of oscillations. The JND is also the amplitude that must be added to the PSE to elicit a quasi-systematic detection ( $PSE + JND =$  quasi-systematic detection threshold).

### 3.1.4. Results

As illustrated in Figure 6, the PSE was associated with an amplitude of trunk oscillation of  $0.57^\circ$  and the JND was equal to an amplitude of  $0.20^\circ$ . Our results indicate that participants were able to detect oscillation from  $0.57^\circ$  of oscillation (detection threshold) and that an amplitude of  $0.77^\circ$  ( $=PSE + JND$ ) of oscillation led to a quasi-systematic detection.

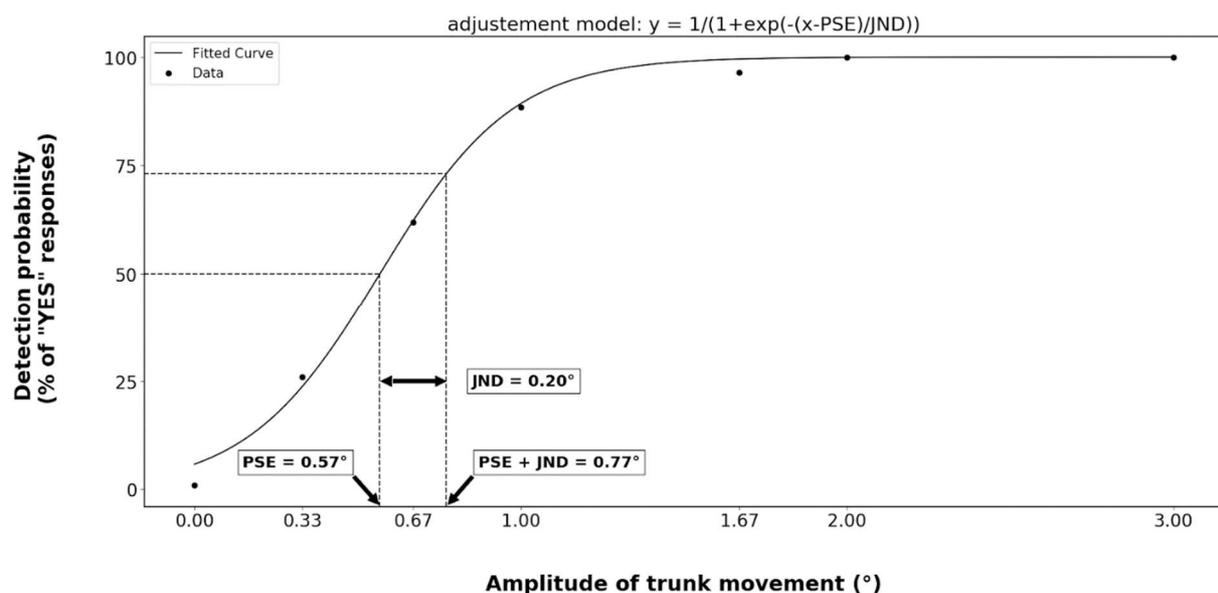


Figure 6 – Proportion of detection according to the amplitude of trunk movement

### **3.2. Summary of Experiment 2**

The results show that trunk movement at an amplitude of  $0.33^\circ$  was not detected by participants (under the  $PSE=0.57^\circ$ ). Participants likely did not perceive the avatar's trunk oscillation in the condition  $0.67^\circ$ , which was situated under the systematic detection threshold ( $PSE + DJP = 0.77^\circ$ ). Moreover, the larger amplitudes of trunk oscillation (from  $1^\circ$  to  $3^\circ$ ) were systematically detected.

Thus, except for the condition  $0.33^\circ$  which cannot be classified as idle motion, these results confirmed the categorization that was suggested in *Experiment 1*, namely that the condition  $0.67^\circ$  may refer to idle motion and the conditions  $1^\circ$ ,  $1.66^\circ$ ,  $2^\circ$ ,  $3^\circ$  may be interpreted as the body expression of a sustained pain when combined with facial pain expression.

## **4. General Discussion**

In the field of affective and social neuroscience, the increasing use of naturalistic, yet well-controlled CG characters (avatars) has led to a new multidisciplinary research avenue working towards the development of anthropomorphic characters, which not only look fully human but also convey 'genuine' emotions. The present work tested whether body motion combined with facial expression increased the perceived intensity and the sense of believability towards the pain expressed by such CG characters, i.e. an avatar from a platform that is specifically dedicated to the study of empathy (Jackson et al., 2015). Moreover, pupil size was recorded during the interaction with the CG character to assess whether pain expression of the avatar elicited a greater emotional arousal when body animation was there.

#### ***4.1. Effect of trunk oscillation as idle motion on participants' perception of the avatar's pain***

Results of Experiment 1 showed that the condition  $0.33^\circ$  did not modulate participants' perception of the avatar's pain, as indicated by the lack of significant difference between this condition and the static condition, either on perceived *Pain Intensity* and *Pain Believability* or on emotional arousal inferred by pupil size. Results of Experiment 2 confirmed that the oscillation at  $0.33^\circ$  was not detected by the participants and, thus, was too low to be considered as idle motion.

By contrast, the condition  $0.67^\circ$  was likely inferred as idle motion since it corresponded to its definition: the oscillation was detectable as it was largely above the detection threshold. However, it was too subtle to be systematically perceived by the participants, even when their attention was focused on the avatar's trunk movement. Importantly, results for *Experiment 1* revealed that the facial pain expression of a CG character was perceived to be more intense and more convincing (i.e. judged more frequently as actually having been experienced) in the presence of this slight trunk oscillation than when the avatar's body was still. It is likely that the avatar looked more 'alive' when it was animated with subtle trunk movement than when it remained still, providing more consistency between graphical and behavioral realism. In fact, the 'statue-like' behavior of the avatar would mismatch with the dynamic facial expression in the static condition. Thus, the facial pain expression would convey less emotion, resulting in a lower perception of pain intensity and believability than in the idle condition.

Reuten and collaborators (2018) have previously shown that the pupil responds less to emotional facial expressions of robots that were judged as uncanny by the participants, compared to more human-like robots. In concordance with this study, it was hypothesized that if the avatar's facial pain expression was judged to be more intense and believable in the idle

condition, it would also evoke a greater pupil response. However, the change in the avatar's pain perception was not accompanied by a greater emotional arousal when low trunk movement was added, as indicated by a similar pupil size between the static ( $0^\circ$ ) and the idle ( $0.67^\circ$ ) conditions. The lack of idle effect on pupil size may result from the very simple, yet quite artificial method used to animate the avatar's trunk. Indeed, postural oscillations came down to a to and fro movement in the anterior-posterior plan (see section 2.1.3.2. *Material and tasks*), while they referred to a much more complex pattern of motion, often spatially represented by an ellipse (Schubert & Kirchner, 2014). Moreover, although postural oscillations are generally associated with respiration, which directly influences postural control (Hunter & Kearney, 1981), breathing movements were not included here. Finally, the avatar's trunk movement was not natural, since it was applied 'mechanically' into the animation software. Motion capture would have been much more suitable as it would take into account natural and unconstrained movements, which are well-known to improve the avatar's acceptability (Pan & Hamilton, 2018; Piwek, McKay & Pollick, 2014, Thompson, Trafton & McKnight, 2011). Moreover, a recent study demonstrated that natural (i.e., biological) motion produced greater pupil dilation than mechanical motion does (William, Cristino & Cross, 2019), which supports the use of natural motion when assessing emotional arousal elicited by a CG character. However, avatar movements with motion capture would need to be pre-recorded and relayed, which implies constraints to a potential use of avatars in an interactive and flexible manner.

Overall, even though the method used for body animation limits the ecological scope of the study, our results showed a large effect size (see *appendix* for Cohen values) regarding the increase of perceived pain intensity and believability in the presence of subtle trunk movement. Idle motion, in particular the movement reproducing human equilibrium, may be

critical when studying emotion perception, as it favors the avatar's anthropomorphism and makes sure that it really conveys emotion.

#### ***4.2. Effect of trunk oscillation as pain behavior on participants' perception of the avatar's pain***

In *Experiment 1*, the significant increase of pain intensity scores from conditions  $1^\circ$  to  $3^\circ$  indicated that the wider the trunk movement was, the more intense the avatar's pain was judged, even in the condition with a voluntary exaggerated oscillation ( $3^\circ$ ). Pain believability rating was also widely improved when the avatar's trunk oscillated from  $1^\circ$  to  $3^\circ$  compared to the static ( $0^\circ$ ) and idle ( $0.67^\circ$ ) conditions. These results support the idea that when systematically detectable, the trunk rocking in this context may evoke a pain behavior. Although this interpretation is yet to be verified, it was proposed that the trunk oscillation might reflect the bodily expression of a sustained pain. Overall, this work is concordant with seminal studies showing that the facial expression of a CG character may convey more emotion when combined with body cues (Buisine et al., 2014; Clavel et al., 2009; De Melo et al., 2010). Moreover, contrary to pain intensity, pain believability did not significantly vary between conditions  $1^\circ$  to  $3^\circ$ , suggesting that it reached its peak at  $1^\circ$  of trunk oscillation. Although partially different than those expected a priori, these results confirmed that the relation between pain believability and perceived pain intensity is not linear. It is not enough to make the CG character express more and more pain to make his pain seem more realistic and convincing.

Importantly, the present study goes even further by revealing that this assumed pain behavior may be easily manipulated to intensify the avatar's facial pain expression, as well as enhancing its accuracy. Indeed, in the static condition, the level of perceived pain in the avatar

was much lower (mean score = 25.21, 95% CI [18.06, 32.36]) than the level of the modeled pain expression (60 % of AU's maxima), while there was a relatively good accuracy when the pain facial expression was combined with a large trunk movement (e.g.  $1^\circ$  = 36.93, 95% CI [30.39, 43.47] ;  $3^\circ$  = mean score 55.38, 95% CI [48.02, 62.74]). A similar underestimation of the avatar's pain intensity has also been observed in a recent study using the EEVEE platform (Tessier and al., 2019). Participants were presented with video clips displaying a sequential or synchronized dynamic facial pain expression of four different avatars modeled at the maxima of pain AUs (100 %). The mean score of pain intensity ranged from 47 to 57 (values from a 0-100 scale) according to the condition, which was half as less as the level of the modeled pain expression. Herein, the presence of trunk oscillation would enrich the avatar's pain expression, resulting in a better congruence between the subjective perception of its pain and the 'presumed' pain level of its facial expression. It would be interesting to explore whether other pain behaviors such as freezing (Defrin, Lotan & Pick, 2006; Feldt, 2000) would induce similar results. If so, this would offer a varied range of pain behaviors while keeping a high-controlled experimental set-up. It would also be interesting to investigate the influence of bodily expression combined with a neutral facial expression to determine whether people can infer pain from bodily expression alone or if facial expression is necessary. Atkinson, Dittrich, Gemmell and Young, 2004, showed that participants were able to detect emotion from points of light representing the shape of actors expressing emotions. We could then suppose that participants would be able to infer pain from body language alone, but this remains to be tested by adding a condition with a neutral expression.

Results for pupil size also led to an unexpected pattern, revealing a significant higher pupil dilation for the conditions  $1^\circ$  and  $3^\circ$ , and no change for the conditions  $1.67^\circ$  and  $2^\circ$  (Figure 5). We proposed that the greater pupil dilation for the condition  $3^\circ$  would rather be an artifact resulting from the 'gap' created between this condition and the other *pain behavior*

conditions. Indeed, the oscillation of  $3^\circ$  would contrast with the other conditions, whose the difference in amplitude was approximately the same. Condition  $3^\circ$  would have been perceived implicitly as an ‘isolated’ event, provoking a greater pupil dilation each time this video was displayed. In other terms, pupil change would not reflect a greater emotional arousal due to the perception of pain in the avatar, but would be rather a response to its exaggerated trunk movement. This alternative explanation is in agreement with a study showing that pupil dilation may reflect not only emotional arousal, but also mechanisms underlying unconscious processing of surprising events (Alamia, Van Rullen, Pasqualotto, Mouraux & Zénon, 2019). Consequently, it is assumed that the pupil responded more to the avatar’s pain expression only when its trunk oscillated at  $1^\circ$  compared with the static condition. The higher pupil response may reflect a higher emotional arousal due to the enhanced realism consistency of the avatar’s pain expression, i.e. between facial and bodily expression of pain. Although speculative, this may support the fact that this trunk amplitude is the optimal rocking level to combine with a facial pain expression modeled from AUs contracted at an intensity of 60 %.

## **5. Conclusion**

This study complements the existing human-machine interaction literature aiming to disentangle how realism consistency between appearance and behavior of a CG character may affect the way we interact with it. Our results showed, for the first time, that both non-communicative and emotional body motion combined with facial expression made the avatar’s pain perception seem more accurate, and in some cases provoked a greater emotional arousal than facial pain expression with a still body. This is in line with the bulk of evidence demonstrating the potential of body language, when congruent with facial expression, to make avatars’ emotions seem more realistic, and thus ‘acceptable’ enough to study underlying processes in humans (Bogdanovych et al., 2016). Another novelty of the present work was to

focus on a pain-expressing CG character. Pain expression is rarely explored in studies questioning avatar realism, which is paradoxical considering the growing number of publications in the medical and pain research domains and the added value of virtual reality for medical training, or the understanding and remediation of pain communication (Brunet-Gouet, Oker, Martin, Grynszpan & Jackson, 2016; Meugnot & Jackson, 2016). Although several limits have been pointed out, this work may encourage researchers to pursue the endeavor of developing fully expressing CG character platform dedicated to cognitive science and health research.

## 6. Appendix - Post-hoc comparison tables

*Table 1 – Pairwise comparisons for the perceived intensity*

		Mean Difference	SE	t	Cohen's d	p <sub>bonf</sub>
0°	0.33°	-1.350	1.110	-1.216	-0.248	1.000
	0.67°	-5.946	1.366	-4.353	-0.889	0.005
	1°	-11.721	2.093	-5.600	-1.143	< .001
	1.67°	-15.296	2.727	-5.610	-1.145	< .001
	2°	-21.788	3.223	-6.760	-1.380	< .001
	3°	-30.175	4.169	-7.237	-1.477	< .001
0.33°	0.67°	-4.596	1.026	-4.481	-0.915	0.004
	1°	-10.371	1.656	-6.262	-1.278	< .001
	1.67°	-13.946	2.235	-6.240	-1.274	< .001
	2°	-20.438	2.770	-7.378	-1.506	< .001
	3°	-28.825	3.711	-7.767	-1.585	< .001
0.67°	1°	-5.775	1.208	-4.780	-0.976	0.002
	1.67°	-9.350	1.943	-4.812	-0.982	0.002
	2°	-15.842	2.442	-6.488	-1.324	< .001
	3°	-24.229	3.465	-6.993	-1.427	< .001
1°	1.67°	-3.575	1.493	-2.395	-0.489	0.528
	2°	-10.067	1.745	-5.770	-1.178	< .001
	3°	-18.454	2.762	-6.682	-1.364	< .001
1.67°	2°	-6.492	1.273	-5.100	-1.041	< .001
	3°	-14.879	2.071	-7.186	-1.467	< .001
2°	3°	-8.387	1.577	-5.319	-1.086	< .001

**Table 2 – Pairwise comparisons for the perceived believability**

		Mean Difference	SE	t	Cohen's d	p <sub>bonf</sub>
0°	0.33°	-6.667	4.156	-1.604	-0.327	1.000
	0.67°	-18.750	5.011	-3.742	-0.764	0.022
	1°	-40.000	5.710	-7.005	-1.430	< .001
	1.67°	-47.083	8.242	-5.713	-1.166	< .001
	2°	-50.833	7.540	-6.742	-1.376	< .001
	3°	-49.583	8.351	-5.937	-1.212	< .001
0.33°	0.67°	-12.083	3.946	-3.062	-0.625	0.116
	1°	-33.333	6.039	-5.519	-1.127	< .001
	1.67°	-40.417	7.260	-5.567	-1.136	< .001
	2°	-44.167	7.296	-6.054	-1.236	< .001
	3°	-42.917	7.601	-5.646	-1.152	< .001
0.67°	1°	-21.250	6.059	-3.507	-0.716	0.040
	1.67°	-28.333	7.388	-3.835	-0.783	0.018
	2°	-32.083	7.372	-4.352	-0.888	0.005
	3°	-30.833	7.636	-4.038	-0.824	0.011
1°	1.67°	-7.083	5.789	-1.224	-0.250	1.000
	2°	-10.833	4.253	-2.547	-0.520	0.378
	3°	-9.583	5.467	-1.753	-0.358	1.000
1.67°	2°	-3.750	5.241	-0.716	-0.146	1.000
	3°	-2.500	5.492	-0.455	-0.093	1.000
2°	3°	1.250	3.422	0.365	0.075	1.000

**Table 3 – T-test relative to 0 for pupil size**

	t	df	p	Cohen's d
0.33°	-0.089	22	0.930	-0.019
0.67°	0.585	22	0.564	0.122
1°	2.654	22	0.014	0.553
1.67°	1.925	22	0.067	0.401
2°	0.848	22	0.405	0.177
3°	2.732	22	0.012	0.570

## 7. Declarations of interest

None.

## References

- Alamia, A., VanRullen, R., Pasqualotto, E., Mouraux, A., & Zenon, A. (2019). Pupil-Linked Arousal Responds to Unconscious Surprisal. *The Journal of Neuroscience*, *39*(27), 5369-5376. <https://doi.org/10.1523/JNEUROSCI.3010-18.2019>
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion Perception from Dynamic and Static Body Expressions in Point-Light and Full-Light Displays. *Perception*, *33*(6), 717-746. <https://doi.org/10.1068/p5096>
- Bogdanovych, A., Trescak, T., & Simoff, S. (2016). What makes virtual agents believable? *Connection Science*, *28*(1), 83-108. <https://doi.org/10.1080/09540091.2015.1130021>
- Bombari, D., Schmid Mast, M., Canadas, E., & Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology : Virtues, pitfalls, and future challenges. *Frontiers in Psychology*, *6*. <https://doi.org/10.3389/fpsyg.2015.00869>
- Bonnet, C. (1986). *Manuel pratique de psychophysique*. A. Colin.
- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602-607. <https://doi.org/10.1111/j.1469-8986.2008.00654.x>
- Brenton, H., Gillies, M., Ballin, D., & Chatting, D. (2005). The Uncanny Valley: does it exist and is it related to presence. *Presence connect*.
- Brunet-Gouet, E., Oker, A., Martin, J.-C., Grynszpan, O., & Jackson, P. L. (2016). Editorial : Advances in Virtual Agents and Affective Computing for the Understanding and Remediation of Social Cognitive Disorders. *Frontiers in Human Neuroscience*, *9*. <https://doi.org/10.3389/fnhum.2015.00697>

- Buisine, S., Courgeon, M., Charles, A., Clavel, C., Martin, J.-C., Tan, N., & Grynszpan, O. (2014). The Role of Body Postures in the Recognition of Emotions in Contextually Rich Scenarios. *International Journal of Human-Computer Interaction*, 30(1), 52-62. <https://doi.org/10.1080/10447318.2013.802200>
- Clavel, C., Plessier, J., Martin, J.-C., Ach, L., & Morel, B. (2009). Combining Facial and Postural Expressions of Emotions in a Virtual Character. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. H. Vilhjálmsson (Éd.), *Intelligent Virtual Agents* (Vol. 5773, p. 287-300). [https://doi.org/10.1007/978-3-642-04380-2\\_31](https://doi.org/10.1007/978-3-642-04380-2_31)
- Courbalay, A., Deroche, T., Prigent, E., Chalabaev, A., & Amorim, M.-A. (2015). Big Five personality traits contribute to prosocial responses to others' pain. *Personality and Individual Differences*, 78, 94-99. <https://doi.org/10.1016/j.paid.2015.01.038>
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249. <https://doi.org/10.1038/nrn1872>
- de Gelder, B., de Borst, A. W., & Watson, R. (2015). The perception of emotion in body expressions : Emotional body perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(2), 149-158. <https://doi.org/10.1002/wcs.1335>
- de Gelder, B., Kätsyri, J., & de Borst, A. W. (2018). Virtual reality and the new psychophysics. *British Journal of Psychology*, 109(3), 421-426. <https://doi.org/10.1111/bjop.12308>
- de Melo, C. M., Kenny, P., & Gratch, J. (2010). Influence of autonomic signals on perception of emotions in embodied agents. *Applied Artificial Intelligence*, 24(6), 494-509.

- Defrin, R., Lotan, M., & Pick, C. G. (2006). The evaluation of acute pain in individuals with cognitive impairment : A differential effect of the level of impairment: *Pain*, *124*(3), 312-320. <https://doi.org/10.1016/j.pain.2006.04.031>
- Egges, A., & Molet, T. (2004). Personalised real-time idle motion synthesis. *12th Pacific Conference on Computer Graphics and Applications, 2004. PG 2004. Proceedings.*, 121-130. <https://doi.org/10.1109/PCCGA.2004.1348342>
- Ekman, P., & Friesen, W. V. (1978). *Facial action coding systems*. Consulting Psychologists Press.
- Feldt, K. S. (2000). The checklist of nonverbal pain indicators (CNPI). *Pain Management Nursing*, *1*(1), 13-21. <https://doi.org/10.1053/jpmn.2000.5831>
- Groom, V., Nass, C., Chen, T., Nielsen, A., Scarborough, J. K., & Robles, E. (2009). Evaluating the effects of behavioral realism in embodied agents. *International Journal of Human-Computer Studies*, *67*(10), 842-849. <https://doi.org/10.1016/j.ijhcs.2009.07.001>
- Hartmann, B., Mancini, M., & Pelachaud, C. (2006). Implementing Expressive Gesture Synthesis for Embodied Conversational Agents. In S. Gibet, N. Courty, & J.-F. Kamp (Éd.), *Gesture in Human-Computer Interaction and Simulation* (p. 188-199). [https://doi.org/10.1007/11678816\\_22](https://doi.org/10.1007/11678816_22)
- Hunter, I. W., & Kearney, R. E. (1981). Respiratory components of human postural sway. *Neuroscience Letters*, *25*(2), 155-159. [https://doi.org/10.1016/0304-3940\(81\)90324-4](https://doi.org/10.1016/0304-3940(81)90324-4)
- Jackson, P. L., Michon, P.-E., Geslin, E., Carignan, M., & Beaudoin, D. (2015). EEVEE : The Empathy-Enhancing Virtual Evolving Environment. *Frontiers in Human Neuroscience*, *9*. <https://doi.org/10.3389/fnhum.2015.00112>

Kätysri, J., Förger, K., Mäkäraänen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in psychology*, 6, 390.

Kätysri, J., de Gelder, B., & Takala, T. (2019). Virtual Faces Evoke Only a Weak Uncanny Valley Effect : An Empirical Investigation With Controlled Virtual Face Images. *Perception*, 48(10), 968-991. <https://doi.org/10.1177/0301006619869134>

Kret, M. E., Roelofs, K., Stekelenburg, J. J., & de Gelder, B. (2013). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Frontiers in Human Neuroscience*, 7. <https://doi.org/10.3389/fnhum.2013.00810>

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, 97(3), 377-395. <https://doi.org/10.1037/0033-295X.97.3.377>

MacDorman, K. F., Green, R. D., Ho, C.-C., & Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior*, 25(3), 695-710. <https://doi.org/10.1016/j.chb.2008.12.026>

Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners : A quantitative cartography of the Uncanny Valley. *Cognition*, 146, 22-32. <https://doi.org/10.1016/j.cognition.2015.09.008>

Meugnot, A., & Jackson, P. L. (2016). The Contribution of New Technological Breakthroughs to the Neuroscientific Research of Pain Communication. In S. van Rysewyk (Éd.), *Meanings of Pain* (p. 87-106). [https://doi.org/10.1007/978-3-319-49022-9\\_6](https://doi.org/10.1007/978-3-319-49022-9_6)

Mitz, A. R., Chacko, R. V., Putnam, P. T., Rudebeck, P. H., & Murray, E. A. (2017). Using pupil size and heart rate to infer affective states during behavioral neurophysiology and

neuropsychology experiments. *Journal of Neuroscience Methods*, 279, 1-12.  
<https://doi.org/10.1016/j.jneumeth.2017.01.004>

Mori, M. (1970/2012). The uncanny valley (K. F. MacDorman & N. Kageki, Trans.). *IEEE Robotics and Automation*, 19(2), 98–100. <http://dx.doi.org/10.1109/MRA.2012.2192811>.

Normoyle, A., Liu, F., Kapadia, M., Badler, N. I., & Jörg, S. (2013). The effect of posture and dynamics on the perception of emotion. *Proceedings of the ACM Symposium on Applied Perception - SAP '13*, 91. <https://doi.org/10.1145/2492494.2492500>

Pan, X., & Hamilton, A. F. de C. (2018). Why and how to use virtual reality to study human social interaction : The challenges of exploring a new research landscape. *British Journal of Psychology*, 109(3), 395-417. <https://doi.org/10.1111/bjop.12290>

Perlin, K. (1995). Real time responsive animation with personality. *IEEE Transactions on Visualization and Computer Graphics*, 1(1), 5-15. <https://doi.org/10.1109/2945.468392>

Piwek, L., McKay, L. S., & Pollick, F. E. (2014). Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition*, 130(3), 271-277.  
<https://doi.org/10.1016/j.cognition.2013.11.001>

Prkachin, K. M. (1992). The consistency of facial expressions of pain: a comparison across modalities. *Pain*, 51(3), 297-306.

Reuten, A., van Dam, M., & Naber, M. (2018). Pupillary Responses to Robotic and Human Emotions : The Uncanny Valley and Media Equation Confirmed. *Frontiers in Psychology*, 9, 774. <https://doi.org/10.3389/fpsyg.2018.00774>

Russell, J. A. (1979). Affective space is bipolar. *Journal of personality and social psychology*, 37(3), 345.

Schubert, P., & Kirchner, M. (2014). Ellipse area calculations and their applicability in posturography. *Gait & Posture*, *39*(1), 518-522. <https://doi.org/10.1016/j.gaitpost.2013.09.001>

Tessier, M.-H., Gingras, C., Robitaille, N., & Jackson, P. L. (2019). Toward dynamic pain expressions in avatars: Perceived realism and pain level of different action unit orders. *Computers in Human Behavior*, *96*, 95-109. <https://doi.org/10.1016/j.chb.2019.02.001>

Thompson, J. C., Trafton, J. G., & McKnight, P. (2011). The Perception of Humanness from the Movements of Synthetic Agents. *Perception*, *40*(6), 695-704. <https://doi.org/10.1068/p6900>

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*(8), 1293-1313. <https://doi.org/10.3758/BF03194544>

Williams, E. H., Cristino, F., & Cross, E. S. (2019). Human body motion captures visual attention and elicits pupillary dilation. *Cognition*, *193*, 104029. <https://doi.org/10.1016/j.cognition.2019.104029>

Wykowska, A., Chaminade, T., & Cheng, G. (2016). Embodied artificial agents for understanding human social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693), 20150375. <https://doi.org/10.1098/rstb.2015.0375>

Wu, Y., Babu, S. V., Armstrong, R., Bertrand, J. W., Luo, J., Roy, T., ... & Fasolino, T. (2014). Effects of virtual human animation on emotion contagion in simulated inter-personal experiences. *IEEE transactions on visualization and computer graphics*, *20*(4), 626-635. <https://doi.org/10.1109/TVCG.2014.19>