



The Role of Body Postures in the Recognition of Emotions in Contextually Rich Scenarios

Stéphanie Buisine , Matthieu Courgeon , Aurélien Charles , Céline Clavel , Jean-Claude Martin , Ning Tan & Ouriel Grynszpan

To cite this article: Stéphanie Buisine , Matthieu Courgeon , Aurélien Charles , Céline Clavel , Jean-Claude Martin , Ning Tan & Ouriel Grynszpan (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, DOI: [10.1080/10447318.2013.802200](https://doi.org/10.1080/10447318.2013.802200)

To link to this article: <https://doi.org/10.1080/10447318.2013.802200>



Accepted author version posted online: 22 Jul 2013.
Published online: 22 Jul 2013.



Submit your article to this journal [↗](#)



Article views: 334



View Crossmark data [↗](#)



Citing articles: 12 View citing articles [↗](#)

The Role of Body Postures in the Recognition of Emotions in Contextually Rich Scenarios

Stéphanie Buisine¹, Matthieu Courgeon², Aurélien Charles^{1,2}, Céline Clavel², Jean-Claude Martin², Ning Tan², and Ouriel Grynszpan³

¹*Arts et Métiers ParisTech, Paris, France*

²*LIMSI-CNRS & Paris-South University, Orsay Cedex, France*

³*Hôpital de la Salpêtrière, Paris, France*

In this article the role of different categories of postures in the detection, recognition, and interpretation of emotion in contextually rich scenarios, including ironic items, is investigated. Animated scenarios are designed with 3D virtual agents in order to test 3 conditions: In the “still” condition, the narrative content was accompanied by emotional facial expressions without any body movements; in the “idle” condition, emotionally neutral body movements were introduced; and in the “congruent” condition, emotional body postures congruent with the character’s facial expressions were displayed. Those conditions were examined by 27 subjects, and their impact on the viewers’ attentional and emotional processes was assessed. The results highlight the importance of the contextual information to emotion recognition and irony interpretation. It is also shown that both idle and emotional postures improve the detection of emotional expressions. Moreover, emotional postures increase the perceived intensity of emotions and the realism of the animations.

1. INTRODUCTION

Affective body expression, perception, and recognition raise many issues such as the respective contributions of the face and the body in affective communication, or the contribution of still postures versus body movement (Kleinsmith & Bianchi-Berthouze, 2012). This study investigates the role of body postures, in addition to verbal utterances and facial expressions, in the processing of ambiguous emotional situations. Several researchers studied body postures associated with basic emotions: Using static images of body postures with emotional information removed from the face, it was

This study was partly supported by a grant from La Fondation de France and La Fondation Adrienne et Pierre Sommer, as part of a project on Virtual Environments for Socio-Cognitive Training in Autism (EVESCA project, Engt n°2007 005874, coordinated by Ouriel Grynszpan).

Address correspondence to Stéphanie Buisine, Arts et Métiers ParisTech, LCPI, 151 bd Hôpital, Paris 75013, France. E-mail: stephanie.buisine@ensam.eu

observed that people accurately recognize basic emotions (De Silva & Bianchi-Berthouze, 2004; Pitterman & Nowicki, 2004; Schouwstra & Hoogstraten, 1995). Results obtained by Coulson (2004) showed that the highest agreement (up to 90% and more) was found for the identification of anger, sadness, and happiness; intermediate agreement for fear and surprise; and the poorest agreement (below 50%) for disgust. According to Coulson, disgust might be primarily communicated through the face. As suggested by the results from Tuminello and Davidson (2011), decoding of bodily expression might be impacted by ethnicity (i.e., higher recognition scores in-group than in cross-ethnic situations) but not by gender stereotyping: Contrary to face decoding, there seems to be no bias toward attributing more masculine emotions to male bodies or feminine emotions to female bodies.

Wallbott (1998) used a different paradigm, with complex emotions, video recordings instead of static pictures, and facial information left available to the viewers. The observers were instructed to code movements and postures of 224 clips representing 14 different emotions. A subsequent discriminant analysis showed that these behavioral categories lead to 54% of correct classification of the intended emotion, but it cannot be excluded that facial expressions influenced the coding of movements and postures. Indeed, Ruffman, Sullivan, and Dittrich (2009) showed that recognition of emotions is significantly higher in a body+face condition than in body-only condition. They also showed that recognition of bodily expressions of basic emotions is still high (58.5%) when only some points of the body are visible (“point-light” condition: head, shoulders, elbows, hands, hips, knee, and feet). The importance of the dynamic nature of emotional expression is also supported by the fact that some emotional patterns (e.g., hot and cold anger) are better recognized from dynamic video clips than from static pictures (Bänziger, Grandjean, & Scherer, 2009). Another example can be taken from the emotions conveyed by body movements in dance that are consistently perceived, even without facial expressions (Sakata, Shiba, Maiya, & Tadenuma, 2004).

Several studies explored the perception of multimodal expressions involving several emotions at the same time (De Gelder & Van den Stock, 2011; Paterson, Pollick, & Jackson, 2002). Meeren, van Heijnsbergen, and de Gelder (2005) tested two combinations of facial and bodily expressions: Using pictures of fearful and angry faces and bodies, they created congruent expressions (face and body associated to the same emotion), and incongruent expressions (e.g., a fearful face on an angry posture). Viewers were instructed to judge the expression of the face, and the results show that congruency significantly increased accuracy of recognition and decreased reaction time. Willis, Palermo, and Burke (2011) investigated the respective contributions of static facial and bodily expressions of emotions on judgments of approachability. They showed that judgments and recognition of incongruent combinations are mainly driven by the facial expression, hence the authors conclude that the meaning of bodily expressions critically depends on the associated facial expressions. Some researchers in the virtual-agents community also explored incongruent combinations of facial and postural expressions of emotions. For example Clavel, Plessier, Martin, Ach, and Morel (2009) used short mute animations lasting 3 s to examine the effects of incongruency, and they observed that recognition predominantly relies on the facial signals of the animated character. Furthermore, they observed that incongruent postures are interpreted in the subjects' judgment as modulations in arousal. This pattern was also confirmed by a more recent study using static pictures of a virtual character displayed from the front or from the side, suggesting that these results might be similar with various viewing angles (Courgeon, Clavel, Tan, & Martin, 2011). Finally, Pollick, Paterson, and Mamassian (2004) showed that body movements could either boost or diminish the effectiveness of facial information in congruent combinations. These studies provide insight on how subjects perceive multimodal expressions and blends of emotions, which occur more frequently in everyday situations than single basic emotions (Abrilian, Devillers, Buisine, & Martin, 2005; Ekman & Friesen, 1975).

The specific contribution of postural idle movements (which are not intended to express any specific emotion) to users' perception of gestures was studied in the field of virtual characters. Luo, Kipp, and Neff (2009) designed a system that combines motion capture data for the lower body with a procedural animation system for arm gestures. A Markov model and association rules were used to predict appropriate postural movements for specified gestures. The authors observed that the use of motion-captured lower body motion was perceived as being more natural than either interpolation of lower body movements or no lower body movement.

The aforementioned studies show the complexity of the perception of emotions when facial and postural expressions are considered in isolation from any contextual information. However, contextual and situational items are likely to influence the way the emotional expressions are processed (Lankes, Bernhaupt, & Tscheligi, 2007, 2010): Whereas viewers mainly

rely on facial expressions when they have to judge static pictures, it was shown that they rely predominantly on contextual information when they have to judge emotional stimuli in interactive scenarios (Lankes & Bernhaupt, 2011). As far as artificial agents are concerned, it was also shown that their appearance influence the interpretation of their attitudes (Komatsu & Yamada, 2011). If contextual information provides a background to emotion recognition, it also introduces a new degree of complexity. For example, the integration of emotional expressions in scenarios raises the issue of the temporal coordination of modalities within the scenario (Buisine, Wang, & Grynszpan, 2010; Yamamoto & Watanabe, 2008). Contextual information also opens up the possibility for even more challenging types of multimodal combinations, such as the expression of irony. Irony is a communicative act in which several meanings are communicated: a literal one, and another one that is contrasting, or even antiphrastic (from Greek *anti* = opposite + *phrasis* = saying), as called in rhetoric (Poggi, Cavicchio, & Magno Caldognetto, 2007). In an ironic act the "true" meaning, the one really intended, is not the one communicated by the literal meaning: It must be understood through inference. According to Poggi et al., one way to alert the addressee to the presence of irony is to use paracommunication: It means that one produces, either simultaneously or in sequence, several signals communicating incompatible meanings. For example, irony can be signaled by a contradiction between the meanings conveyed in words (the literal meaning) and in nonverbal communication. In such a situation, the hearer has to compare the two opposite meanings and decide which of them, on the basis of the context, is more plausible.

In the present study, we intend to investigate such complex cases of multimodal combinations of emotional expressions. Beyond static pictures or short animations, we are interested in assessing the benefits of postures for a full-body animated character embedded in 3D scenes, who is telling a story that includes ironic and narrative items. Furthermore, we wish to investigate in such scenarios the impact of idle postures, whose role on recognition of emotions is still an issue.

For that purpose we implemented the following kinds of body postures: (a) idle postures, which are natural neutral movements such as small variations in body postures or balance changing (Egges, Molet, & Magnenat-Thalmann, 2004); (b) emotional postures, which are directly and explicitly related to a mental state and cover a much wider range of muscular tension and relaxation; and (c) still postures, which correspond to the absence of body movement and serve as a control condition. These three types of body postures are scaled in terms of communicativeness: The still posture condition is the least demanding in terms of attention; the idle condition will presumably attract the user's attention though to a lesser extent than the emotional condition, which is more demonstrative. On a theoretical viewpoint, this study is expected to help understanding the role of body postures with regard to cognitive and emotional processing. In particular, we consider two opposite hypotheses

in the scope of our experiment: Either postures are distractors that hinder processing of the virtual character's facial expressions, or they provide additional cues that enable a better grasp of the character's message. If the former hypothesis is correct, then the absence of body movements (still condition) should yield the highest performances in emotion recognition, whereas the emotional postures would yield the worst. In the case of the latter hypothesis, the emotional postures would yield the best performances and the still condition the worst. In addition, the still condition produces a rigid posture that could be perceived as strange by the viewer, thus inducing an "uncanny valley" type of phenomenon (Groom et al., 2009).

2. METHOD

2.1. Experimental Material

The material was composed of ironic scenarios, each one containing three emotions (see Table 1), and consequently three different facial expressions and three different postures. The three scenarios were generated three times each (i.e., nine animations in total): once with facial expressions and still posture (still condition), once with facial expressions and idle postures (idle condition), and once with facial expressions and emotional postures (congruent condition). The facial expressions remained constant across the three conditions, and the postures we used are presented in Figure 1. For the third scenario, which contained two instances of anger, we used two different angry postures (see Figure 1).

In each scenario a single virtual character addressed the participant. Scenarios followed a pattern that required integrating both verbal and nonverbal channels in order to build the social context. Although the character was reporting a situation he had been involved in, he uttered a sentence that could be interpreted in two distinct ways according to the context (Table 1). The character displayed facial expressions that enabled disambiguating this key sentence and therefore understanding the whole message. For example an utterance like "Thank you so much" (with a neutral voice intonation) associated to a

facial expression of *anger* should not be interpreted literally as meaning that the character is pleased but rather nonliterally as meaning that the character is being ironical.

Acapela (www.acapela-group.com) was used to generate speech, the experiment being run in French. We used the MARC platform (Courgeon, Martin, & Jacquemin, 2008) to generate real-time animations and control virtual characters in a 3D environment. Virtual characters faces are composed of nearly 20,000 triangles, which enable us to design subtle facial expressions. Facial animation is performed using the MPEG4 keypoint approach, combined with realistic expressive wrinkles rendering. The body animation is achieved using skeleton and rigging based animation. We use a set of 71 bones, including articulated fingers. Every facial keypoint and skeleton joint can be precisely controlled using dedicated 3D editors. The resulting animations are rendered using graphics processing unit-based (GPU-based) animation and rendering techniques.

Facial expressions were specified using Ekman's guidelines (Ekman & Friesen, 1975) as well as the Cambridge Mindreading database (Golan, Baron-Cohen, & Hill, 2006). We selected and specified prototypical postures for the five emotions of interest (sadness, fear, surprise, disgust, and anger) as well as idle postures based on the AffectMe database (Kleinsmith, Rebai, Berthouze, & Martin, 2009) and previous work (Clavel et al., 2009). These modalities were all synchronized in the MARC platform.

We followed a two-step process to validate our experimental material. The first step focused on the facial expressions, which were evaluated by 53 participants (18 women, 35 men; M age = 31.6 years). Static expressions of the five emotions involved in our scenarios (surprise, anger, sadness, disgust, and fear) were displayed in a random order, three times each, and submitted to a classical forced-choice paradigm (one label to be selected out of seven: neutral + six basic emotions). The recognition rates were 98.11% for sadness, 84.91% for anger, 83.65% for surprise, 57.86% for disgust, and 40.24% for fear. Our results are globally consistent with previous findings with static pictures of human facial expressions (Russell, 1994) with high recognition rates for sadness, anger, surprise, and

TABLE 1

The Three Ironic Scenarios, With Utterances Holding a Nonliteral Meaning (in Italics), and Emotional Expressions (in Brackets) Associated With the Previous Utterance

Scenario 1	"I have just realized that the cottage cheese I bought is out of date [surprise]. <i>The cheese cake that I wanted to cook would have been really delicious</i> [disgust]. I will point it out to the shop manager. This is unacceptable [anger]."
Scenario 2	"Our teacher is organizing a boxing tournament today [sadness]. He chose the participants by casting lots: My opponent has been boxing for 10 years [surprise]. <i>How lucky I am to be able to fight against him</i> [fear]."
Scenario 3	"I was not even invited to my best friend's birthday party [sadness]. <i>That is so nice of him</i> [anger]. If this is so, I will not invite him to my birthday party either. We will see how he feels [anger]."

Note. The text in this table is a translation of the scenarios written in French that were actually used during the experiment.



FIG. 1. Emotional and idle postures used in our animations. *Note.* Here they are all presented with a neutral facial expression (color figure available online).

some confusion between disgust and anger. We also observed confusions between fear and surprise, resulting in a recognition rate for fear that is lower than expected. However, similar rates were obtained using virtual faces based on acted expressions (Wallraven, Breidt, Cunningham, & Bülthoff, 2005). Our fear expression was possibly not intense enough, but we judged that an expression of panic fear, including a more stretched mouth and more tensed eyebrows, would not be relevant to

our scenarios. This confusion between fear and surprise was also observed in other studies (Rapcsak et al., 2000) and might be related to common components in the appraisal profiles of surprise and fear (Scherer, 2001).

The second validation step focused on our posture design. We conducted a test with 21 participants (14 men, 7 women; M age = 27.4 years) using short silent animations of the body postures. The character's face was covered with an opaque oval patch to prevent a bias due to the facial expressions' influence on the decoding of postures. The postures were presented in a random order and submitted to a forced-choice paradigm including the six basic emotions and a neutral option. The participants could play the animation as many times as necessary. The recognition scores are detailed in Table 2.

Postural expressions of anger, sadness, and surprise were recognized above the chance level (14.3%). Idle postures were consistently perceived as conveying a neutral state (46.9%), which is very relevant for our research goal. Disgust was most frequently recognized as fear (76.2%) and fear often interpreted as neutral (28.6%). The overall recognition score for our postures is 52.3%.

These results are partly in line with other studies about the perception of postural expressions of emotions. Such studies are fewer than studies about facial expressions, and they are also difficult to compare because they use different sets of emotions and different types of data (e.g., acted postures, dance). For example, recognition rates of 56% were observed for dance motions (Camurri, Lagerlof, & Volpe, 2003), 30% for postural acted emotions superimposed on drinking and knocking actions (Paterson, Pollick, & Sanford, 2001), 67% agreement between observers of naturalistic postures (Kleinsmith, Bianchi-Berthouze, & Steed, 2011), or 55% for 12 basic and complex acted emotions (Dael, Mortillaro, & Scherer, 2012). Similarly to our results, Coulson (2004) reported low recognition scores for postural expressions of disgust.

With these scores in mind regarding the recognition of our facial and postural expressions taken in isolation, the next experiment shows the influence of rich contextual and multimodal information on the recognition of emotions.

2.2. Experimental Procedure

The main experiment with contextually rich scenarios involved 27 participants (16 men, 11 women; M age = 25.2 years, $SD = 3.7$). None of them participated in the preliminary tests. The order of the scenarios and posture conditions (still, idle, and congruent conditions) were counterbalanced across participants. We adopted a between-subject design to compare the three conditions (still, idle, and congruent conditions) for a given scenario to avoid a practice effect in the interpretation of a scenario and to ensure more spontaneous responses. The arrangement of scenarios and conditions was determined by a Latin-square design (see Table 3), so that each participant was exposed to every scenarios and every conditions. The participants watched the animations on a computer screen and had to

TABLE 2
Distribution of Recognition Rates Associated With Our Posture Stimuli (%)

		Response						
		Anger	Disgust	Happiness	Neutral	Fear	Surprise	Sadness
Stimulus	Anger	76.7	4.7	0.0	9.3	2.3	7.0	0.0
	Disgust	0.0	14.3	0.0	4.8	76.2	4.8	0.0
	Idle	12.5	9.4	3.1	46.9	9.4	17.2	1.6
	Fear	0.0	14.3	0.0	28.6	28.6	23.8	4.8
	Surprise	9.5	0.0	0.0	4.8	19.0	66.7	0.0
	Sadness	2.4	16.7	2.4	11.9	2.4	4.8	59.5

Note. Shaded cells indicate the combinations that are expected to produce the highest recognition scores, and bold indicates the actual highest recognition scores. Chance level (1/7) amounts to 14.3%.

TABLE 3
Latin-Square Counterbalancing the Scenarios, Conditions, and Order for the 27 Participants

	Still	Idle	Congruent
Users 1 to 3	Scenario 1	Scenario 2	Scenario 3
Users 4 to 6	Scenario 2	Scenario 3	Scenario 1
Users 7 to 9	Scenario 3	Scenario 1	Scenario 2
	Idle	Congruent	Still
Users 10 to 12	Scenario 1	Scenario 2	Scenario 3
Users 13 to 15	Scenario 2	Scenario 3	Scenario 1
Users 16 to 18	Scenario 3	Scenario 1	Scenario 2
	Congruent	Still	Idle
Users 19 to 21	Scenario 1	Scenario 2	Scenario 3
Users 22 to 24	Scenario 2	Scenario 3	Scenario 1
Users 25 to 27	Scenario 3	Scenario 1	Scenario 2

fill in a questionnaire after each animation, with the following items:

- Number of emotions detected in the animation: Because we had encoded three emotions, we used a 0-to-5 numerical scale for this item to leave room for either overestimation or underestimation of the number of emotions.
- Label of the emotions detected: We used a free-response format so that the answers would not be influenced by a list of response options (Russell, 1994). In particular, we were interested to know whether the emotional expressions involved in ironical utterances would be interpreted as basic or complex emotions, or a combination of both.
- Causes of the emotions expressed by the character: We also used a free-response format for this item, with the aim of capturing whether the participants understood the ironic item in the scenario.

- Level of emotional intensity expressed by the character throughout the animation: We used a Likert-type item, which is an ordinal response format contrasting two semantic ends along a continuum (very low intensity, very high intensity). We introduced 7 scale points in this response format, as recommended by Carifio and Perle (2007).
- Level of realism of the animation with a 7-point Likert-type response format.

To collect more spontaneous data and to limit the duration of the experiment, each animation could be viewed only once. The whole experiment lasted about 15 min per participant.

3. RESULTS

The data set was analyzed by means of analyses of variances with the condition (still, idle, congruent) as the within-subject factor. Fisher's Least Significant Difference was used for post hoc tests. All the analyses were performed with SPSS v.18 (Statistical Program for Social Sciences).

We observed a main effect of the condition on the number of emotions perceived, $F(2, 52) = 3.40, p = .041$, partial $\eta^2 = 0.116$; see Figure 2 left panel). Post hoc comparisons suggested that participants tended to detect fewer emotions in the still condition than in the idle ($p = .083$) and the congruent condition ($p = .031$). There was no significant difference between the idle and congruent conditions. Means and standard errors are detailed in Table 4.

The correct recognition variable was derived from the emotional categories labeled by our participants. Because of the free response format, there was not always a direct mapping between their answers and the emotional labels we expected. For this reason we recruited two independent judges whose role was to decide how many expressions were correctly recognized in each animation. Each scenario was treated as a whole, and the judges had to attribute a global recognition score based on the emotions listed by the participants. Recognition score ranged from 0 to 3 because there were three emotions encoded in each scenario. For example, a response like "sick, ironic, angry, furious"

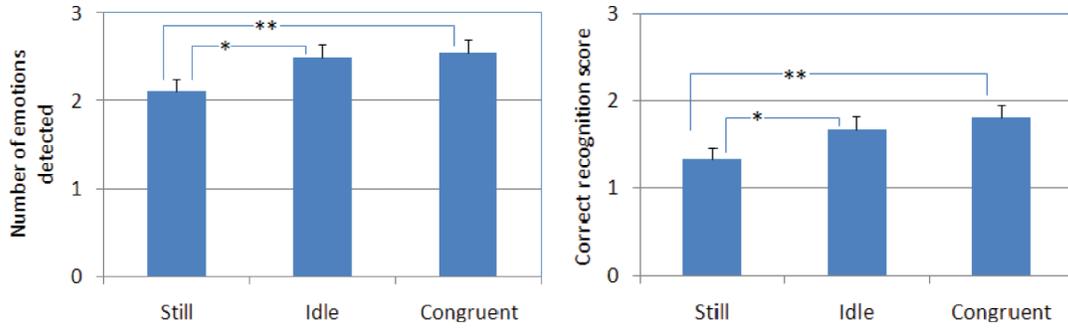


FIG. 2. Effects of the condition (still postures, idle postures, congruent postures) on the number of emotions detected in the animations (left panel) and on the recognition accuracy (right panel). * $p < .1$. ** $p < .05$ (color figure available online).

TABLE 4
Means and Standard Errors for All Condition \times Dependent Variable Combination

Dependent Variable	Still Condition		Idle Condition		Congruent Condition	
	M	SE	M	SE	M	SE
No. of emotions detected	2.111	0.138	2.491	0.150	2.554	0.150
Correct recognition score	1.333	0.131	1.667	0.160	1.815	0.131
Intensity	3.105	0.285	3.474	0.222	4.216	0.231
Realism	3.452	0.317	3.716	0.237	4.293	0.259

to Scenario 1 was attributed a 2-point recognition score: The judges decided that disgust (sick) and anger (angry, furious) were recognized, surprise was not detected, and irony not considered as an emotion. The whole corpus of emotional labels was examined by the two judges who were blind to the condition factor (whether each answer was associated to the still, idle, or congruent condition). Interjudge agreement (Cronbach's alpha) amounted to 0.934 for the recognition scores. In this data set an analysis of variance revealed a main effect of the condition, $F(2, 52) = 3.36$, $p = .042$, partial $\eta^2 = 0.115$, see Figure 2 right panel): Recognition was poorest in the still condition, marginally better in the idle condition ($p = .095$), and significantly higher in the congruent condition ($p = .016$). There was no significant difference between the recognition scores in the idle and congruent conditions. Note should be taken that these results are coherent with the above analysis concerning the number of emotions detected.

Table 5 details the recognition scores for each emotion in each condition and lists the labels that were proposed by the participants and judged as incorrect.

The independent judges also analyzed qualitatively the answers related to the understanding of the ironic scenario. They globally scored each answer as *right* (1) or *wrong* (0) with regard to the contextual information emphasized in the participants' answer. For example a correct answer for Scenario 2 was "about to fight against too-strong a boxer" and an incorrect answer "fight against a valorous opponent." In this respect they considered that the participants caught the nonliteral meaning

of the scenario in 73 trials out of 81 (i.e., 90% of cases). Furthermore, we found 12 times the word "irony" in the participants' emotional labels (see Table 5), which denote a correct interpretation. The misunderstandings (eight of 81 trials) all occurred for Scenario 2 (the boxing tournament): Speech was interpreted literally ("How lucky I am to fight against him") and associated to pleasantness or pride, despite the facial expression of fear. These misunderstandings equally occurred in the still, idle, and congruent conditions.

Regarding the perceived emotional intensity, we found a significant influence of the condition, $F(2, 52) = 3.48$, $p = .038$, partial $\eta^2 = 0.118$, see Figure 3 left panel): A given scenario was judged as more intense in the congruent than in the still condition ($p = .036$) and the idle condition ($p = .015$). There was no significant difference between still and idle conditions.

The condition tended to influence the scores of realism as well, $F(2, 52) = 2.85$, $p = .067$, partial $\eta^2 = 0.099$; see Figure 3 right panel) Congruent postures conveyed more realism than still ($p = .021$) and idle ($p = .059$) postures. There was no significant difference between the control and idle animations.

Finally we computed the correlation matrix (see Table 6) of our four main dependent variables (number of emotions detected, recognition score, intensity, realism) to better understand their mutual links. The results show that intensity and realism are strongly correlated ($r = .66$), detection and recognition of emotions are moderately correlated ($r = .46$), and intensity and recognition are moderately correlated ($r = .44$). The other correlations are weak.

TABLE 5
Percentage of Correct Labeling for Each Emotion in Each Condition, and Incorrect Labels Proposed by the Participants With Their Number of Occurrence in Parentheses (%)

		Still Condition	Idle Condition	Congruent Condition
Scenario 1	Surprise	11.1	55.6	33.3
	Disgust	22.2	33.3	44.4
	Anger	88.9	88.9	100
	Incorrect labels	Disappointment (3), irony (2), frustration, sadness	Disappointment (2), irony (2), depression, regret, sadness	Irony (5), frustration, disappointment, resentment, indignation
Scenario 2	Sadness	44.4	33.3	44.4
	Surprise	22.2	33.3	44.4
	Fear	44.4	44.4	66.7
	Incorrect labels	Pleasantness (2), pride (2), disgust, doubt, irony, impatience	Stress, anger, frustration, pleasantness (2), pride, uneasiness	Skepticism, enjoyment, desire, enthusiasm, anger, disappointment, joy (2), impatience
Scenario 3	Sadness	88.9	88.9	100
	Anger	66.7	100	100
	Incorrect labels	Disgust, resentment, disappointment (4), irony, sulk, grudge	Irony, resentment (2), revenge (2), incomprehension	Revenge (2), resentment (2)

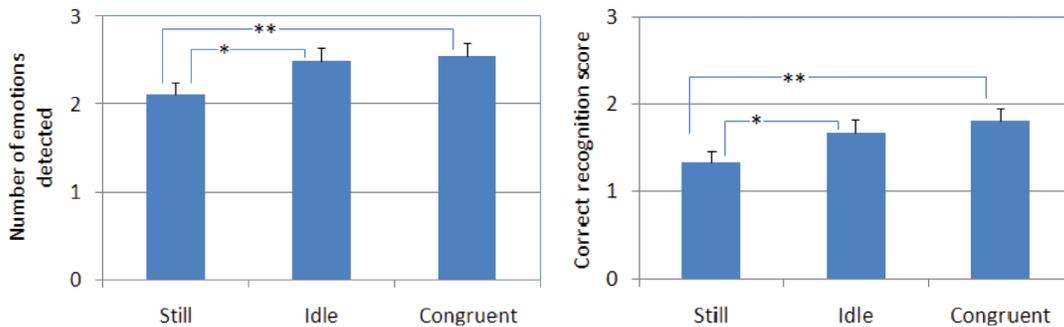


FIG. 3. Effects of the condition (still postures, idle postures, congruent postures) on the perceived intensity of emotions throughout the scenario (left panel) and on the scores of realism (right panel). * $p < .1$. ** $p < .05$ (color figure available online).

TABLE 6
Correlation Matrix Between Our Dependent Variables

	No. of Emotions Detected	Correct Recognition Score	Emotional Intensity	Realism
No. of emotions detected	1	0.46	0.38	0.18
Correct recognition score		1	0.44	0.22
Emotional intensity			1	0.66
Realism				1

4. DISCUSSION

The originality of our study lies in two main aspects. Firstly, contrary to most of the existing literature about body posture, we investigated their role in recognition of emotion in

contextually rich scenarios. Those scenarios, in addition to providing a contextual background to the emotional expressions, implemented a particularly challenging communication situation, namely, the use of irony. Second, beyond the two most

classical conditions investigated in research on body posture (no posture/emotional posture), we introduced a third condition (idle posture), with a role in recognition of emotions that was never studied, to our knowledge. The experiment reported here sought to determine which of two opposite hypotheses was correct, that is, that (a) body postures in animated characters tend to distract attention from the facial expressions or (b) that they provide additional emotional guidance. The results clearly supported the latter one.

Our study provides new unexpected prospects about idle postures and their potential impact on emotion processing. We observed that idle postures, although unrelated to the emotion expressed by the character, appeared to marginally improve emotion detection and recognition with comparison to the still condition. Further experiments are necessary to understand whether these marginal improvements are an experimental artifact related to our particular study and material, or would give rise to robust significant improvements with, for example, a larger user sample. With regard to the attentional and interpretative mechanisms, these results suggest that postures, in general, provide additional cues to the facial expressions and help (rather than distract) the viewer in decoding nonverbal emotional message. The correlation we observed between detection and recognition of emotions also supports the attentional prompting hypothesis: Body postures seem to draw the viewer's attention to the emotional signals in the character's nonverbal behavior. This positive attentional effect might particularly be true for full-body displays of virtual characters and might not be so useful when the relative size of the face is bigger and when facial expressions are easier to notice and decode. Our experiment and results suggest that idle postures could constitute a convenient trade-off for increasing the characters' expressiveness at a lower cost than designing emotional postures. Indeed, the same set of idle postures can be combined to several facial expressions of emotions and still appear to help interpret the character's emotions.

Emotionally congruent body postures nonetheless keep several advantages with comparison to idle postures: They are more effective to support emotion detection and recognition, and they make the emotions more intense in the viewer's perception, which is consistent with Ekman's view that bodily movements provide information on the intensity of emotion (Ekman & Friesen, 1974). Emotionally congruent postures also increased the perceived realism of the virtual characters. Realism is a qualitative feature that is usually sought in the design of virtual emotional characters. It also refers to believability, which is often considered to rely on the visual properties of the character and on the generation of verbal and nonverbal behavior during interaction with the user (Johnson, Rickel, & Lester, 2000). In our study, emotional postures proved to increase both realism and communication effectiveness (detection and recognition of emotions). This is a very positive result in comparison to other research suggesting that realism does not always correlate with communication effectiveness (Buisine et al., 2010; Calder et al.,

2000). For example Calder et al. (2000) showed that caricaturing facial expressions, although decreasing ratings of human likeness or plausibility, may increase recognition of the agents' emotions (shorter reaction times), increase the neural response of the viewer, and increase her or his ratings of emotional intensity. In a study about the temporal relations between speech and facial expressions of emotion, Buisine et al. (2010) also observed that displaying the facial expression at the end of the spoken utterance was the most efficient pattern for maximizing the recognition of the emotion, although it was judged as unrealistic by the same viewers. In contrast in the present study, emotional body postures, because they increased emotional intensity, increased both recognition accuracy (moderate correlation between intensity and recognition score) and realism (strong correlation between intensity and realism). Therefore, with emotional body postures, a trade-off between recognition effectiveness and realism is not necessary because they favor these two parallel goals.

With regard to the recognition of emotions in contextually rich scenarios, our results highlight the importance of the scenario beyond the design of accurate emotional expressions: As can be seen in Table 5, the same facial and postural expressions were not recognized to the same extent according to the scenario they were inserted in. For example, the recognition of sadness in congruent condition, although reaching 100% in Scenario 3, fell to only 44.4% in Scenario 2. Such a pattern resembles those found by Lankes and Bernhaupt (2011) with facial expressions and consequently complement the results obtained in recognition of emotional expressions in isolation from any context: For example, the remarkable capacity of Sadness to be conveyed through body posture alone (Coulson, 2004) when shown in isolation could actually represent a maximum that is likely to be modulated by additional contextual information. In a similar way, the understanding of irony proved to be dependent primarily on the contextual information (i.e., the scenario) rather than on the recognition of emotional nonverbal expressions. Irony was understood in 90% of cases, although sometimes the target emotion meant to alert the viewer about the presence of irony was poorly recognized (e.g., the recognition score of disgust in Scenario 1, still condition, was only 22.2%). The successive verbal utterances were contrasting enough with each other to imply irony—except for Scenario 2, which yielded to some misunderstandings. In the latter case irony was supposed to be signaled by a contradiction between the literal meaning (“How lucky I am”) and the nonverbal expressions of fear. However, as can be seen in our preliminary tests, our fear expressions had low recognition rates, either in facial expression (40.2%) or in posture (28.6%). This may explain why irony was not so well understood in Scenario 2 as in the other scenarios.

5. EXAMPLE APPLICATION FRAMEWORK

The outcome of this could prove useful for developing multimedia tools dedicated for the sociocognitive training of

individuals with high-functioning autism. Autism is defined as a pervasive developmental disorder (American Psychiatric Association, 1994) involving qualitative alterations in social interaction, verbal, and nonverbal communication. High-functioning autism refers to the subgroup that has average or above-average IQ. An increasingly intensive research is dedicated to computer-supported tools that are developed to evaluate and take care of cognitive impairments (Grynszpan, Weiss, Perez-Diaz, & Gal, in press). A trend focuses specifically on the use of virtual characters that provide support for socio-emotional education in autism (Parsons & Cobb, 2011). For instance, Tartaro and Cassell (2006) designed a virtual character resembling a child for training collaborative storytelling skills; Mitchell, Parsons, and Leonard (2007) designed an intervention based on a socially challenging situations in a virtual café; Moore, Cheng, McGrath, and Powell (2005) tested how individuals with high-functioning autism could associate an emotionally connoted situation with the appropriate facial expression of a virtual character; and Schwartz, Bente, Gawronski, Schilbach, and Vogeley (2009) employed virtual humans for evaluating the processing of gaze and facial expressions by individuals with high-functioning autism.

Despite preserved intellectual abilities, individuals with high-functioning autism are reported to experience profound difficulties with irony (Happé, 1993). The ability to integrate different components of expressive behaviors in a timely manner is considered linked to their misunderstanding of social subtleties (Klin, Jones, Schultz, Volkmar, & Cohen, 2002). Emotional nonverbal behaviors seem to be used inappropriately by individuals with autism (Grynszpan, Martin, & Nadel, 2008). In an experiment by Grynszpan and colleagues (Grynszpan et al., 2011; Grynszpan et al., 2012) using eye-tracking, participants with high-functioning autism and typical controls were administered a task where they had to interpret the message of a realistic virtual human character whose verbal utterances and related facial expressions were specifically designed to be contrasting, thus conveying a nonliteral meaning similar to irony. Results showed that participants with autism had significantly lower interpretation scores and spent significantly more time looking at areas outside the face. Considering the diminished attention to faces in autism, the hypothesis that body postures would distract their attention rather than provide them with emotional guidance seems plausible. Therefore, repeating the present study with participants having high-functioning autism could conceivably yield the opposite pattern of results to the one found here. The protocol designed in here seems highly relevant for assessing the ability of people with autism in integrating verbal and nonverbal emotional cues. The emotional detection performances that we collected here in a sample of people without autism can serve as a reference base for future evaluation of individuals with autism on the same task. Our long-term goal is to design a virtual environment for training social skills based on 3D embodied agents that can talk while displaying emotional facial expressions and body movements. In future

research, social embodied agents could be used in conjunction with physiology-based affect recognition systems that monitor children with autism (Liu, Conn, Sarkar, & Stone, 2008). Although we illustrate here a possible application for autism, such expressive embodied virtual agents are highly relevant for training social competencies in other neurodevelopmental disorders. In addition, they could potentially be used to enhance teaching of foreign languages by complementing speech with culturally specific expressions and body movements.

6. GENERAL CONCLUSION

As previously emphasized, the present study provided new original results regarding the decoding of contextually rich, ironical scenarios. We showed how contextual information is likely to influence the decoding of emotional expressions, a result that complements the existing literature about facial and postural information. We also provided an example of implementation of irony, by means of contextual information and paracomunicative items (contrast between verbal literal meaning and nonverbal emotional expression). This kind of subtle communication style, very specific to humans, still represents a challenge for virtual characters in search for believability. Our design of irony proved to be efficient because it was correctly interpreted in 90% of cases. In this respect, we may recommend that animation designers should choose verbal contextual items very carefully, because in our experiment it seemed at least as important as the accuracy of paracomunicative items. An alternate way to alert to irony could be to use metacommunication instead of paracomunication, for example, with specific facial signals like the ironic smile (raising the corner of only one lip), which specifically means "I am being ironic" (Poggi et al., 2007).

We also provided original results about the potential usefulness of idle postures for favoring the decoding of emotions. Even if they are unspecific to any emotion, idle postures seem to improve the recognition of the character's emotional facial signals. Therefore they appear particularly cost-effective because they can be used in combination to various facial expressions of emotions. However, we still recommend that prospective designers use idle postures with caution, as in our experiment they failed to improve realism with comparison to the still condition. This result raises concern and calls for further experimental investigation.

Finally, several limitations of this study draw avenues for future research. First, we conducted this study with a limited sample of users, all young and with a wide experience of new technologies. In addition to the application for high-functioning autism, future research should extend our findings with other kinds of users, for example, children or elderly people whose cognitive processes are either in development or in decline. Such populations will not necessarily be subject to attentional and interpretative effects in the same way as our users were. A second shortcoming of our study is its time frame.

The decoding of emotional information and interpretation of specific behaviors such as irony might depend on the history of the relationship between the speaker and the addressee. Although our experiment was an attempt to provide rich contextual background to the recognition of emotions, it is still far from simulating everyday communication situations. In this respect, longitudinal research should investigate the evolution of communication between virtual characters and human users over longer periods: Along the history of the relationship, does the communication accuracy improve? Do human users get used to the virtual character's communication style and manage to disregard the realism of its behaviors? The contribution of facial expressions versus body movements could also depend on the task (Baylor & Kim, 2009). Finally, future research should also examine the interpretation of irony in situations involving the addressee more personally. Indeed most of the time irony has an evaluative import supposed to touch the addressee, like ironic praise or ironic criticism (Poggi et al., 2007). Again, such investigation requires building longer relationship between virtual characters and human users. Despite the limitations of our study, we believe that it provided new knowledge on the processing of virtual characters' postures and emotional information and will contribute to the long-term goal of designing more believable and more effective autonomous virtual characters.

REFERENCES

- Abrikan, S., Devillers, L., Buisine, S., & Martin, J. C. (2005). EmoTV1: Annotation of real-life emotions for the specification of multimodal affective interfaces. *HCI'05 Human Computer Interaction International*.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Bänziger, T., Grandjean, D., & Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: The Multimodal Emotion Recognition Test (MERT). *Emotion*, 9, 691–704.
- Baylor, A. L., & Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior*, 25, 450–457.
- Buisine, S., Wang, Y., & Grynszpan, O. (2010). Empirical investigation of the temporal relations between speech and facial expressions of emotion. *Journal on Multimodal User Interfaces*, 3, 263–270.
- Calder, A. J., Rowland, D., Young, A. W., Nimmo-Smith, I., Keane, J., & Perrett, D. I. (2000). Caricaturing facial expressions. *Cognition*, 76, 105–146.
- Camurri, A., Lagerlof, I., & Volpe, G. (2003). Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies*, 59, 213–225.
- Carifio, J., & Perle, R.J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, 3, 106–116.
- Clavel, C., Plessier, J., Martin, J.C., Ach, L., & Morel, B. (2009). Combining facial and postural expressions of emotions in a virtual character. *IVA: International Conference on Intelligent Virtual Agents*, 287–300.
- Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, 28, 117–139.
- Courgeon, M., Clavel, C., Tan, N., & Martin, J. C. (2011). Front view vs. side view of facial and postural expressions of emotions in a virtual character. *Journal Transactions on Edutainment*, 6, 132–143.
- Courgeon, M., Martin, J. C., & Jacquemin, C. (2008). MARC: A Multimodal Affective and Reactive Character. *AFFINE: International Workshop on Affective Interaction in Natural Environments*.
- Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion*, 12, 1085–1101.
- De Gelder, B., & Van den Stock, J. (2011). Real faces, real emotions: Perceiving facial expressions in naturalistic contexts of voices, bodies and scenes. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 535–550). Oxford, UK: Oxford University Press.
- De Silva, P., & Bianchi-Berthouze, N. (2004). Modeling human affective postures: An information theoretic characterization of posture features. *Computer Animation and Virtual Worlds*, 15, 269–276.
- Egges, A., Molet, T., & Magnenat-Thalmann, N. (2004). Personalised real-time idle motion synthesis. *Pacific Graphics Conference*, 121–130.
- Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology*, 29, 288–298.
- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face. A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall.
- Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge Mindreading (CAM) Face-voice Battery. *Journal of Autism and Developmental Disorders*, 36, 169–183.
- 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, 842–849.
- Grynszpan, O., Martin, J. C., & Nadel, J. (2008). Multimedia interfaces for users with high functioning autism: An empirical investigation. *International Journal of Human-Computer Studies*, 66, 628–639.
- Grynszpan, O., Nadel, J., Constant, J., Le Barillier, F., Carbonell, N., Simonin, J., & Martin, J. C. (2011). A new virtual environment paradigm for high-functioning autism intended to help attentional disengagement in a social context. *Journal of Physical Therapy Education*, 25, 42–47.
- Grynszpan, O., Nadel, J., Martin, J. C., Simonin, J., Bailleul, P., Wang, Y., . . . Constant, J. (2012). Self-monitoring of gaze in high functioning autism. *Journal of Autism and Developmental Disorders*, 42, 1642–1650.
- Grynszpan, O., Weiss, P. L., Perez-Diaz, F., & Gal, E. (in press). Innovative technology based interventions for Autism Spectrum Disorders: A meta-analysis. *Autism*.
- Happé, F. G. E. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, 48, 101–111.
- Johnson, W. L., Rickel, J., & Lester, J. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, 11, 47–78.
- Kleinsmith, A., & Bianchi-Berthouze, N. (2012). Affective body expression perception and recognition: A survey. *IEEE Transactions on Affective Computing*, 1–20.
- Kleinsmith, A., Bianchi-Berthouze, N., & Steed, A. (2011). Automatic recognition of nonacted affective postures. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 99, 1–12.
- Kleinsmith, A., Rebai, I., Berthouze, N., & Martin, J. C. (2009). Postural expressions of emotion in motion captured database and in a humanoid robot. *AFFINE International Workshop on Affective-aware Virtual Agents and Social Robots*.
- Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Defining and quantifying the social phenotype in autism. *The American Journal of Psychiatry*, 159, 895–908.
- Komatsu, T., & Yamada, S. (2011). How does the agents' appearance affect users' interpretation of the agents' attitudes: Experimental investigation on expressing the same artificial sounds from agents with different appearances. *International Journal of Human-Computer Interaction*, 27, 260–279.
- Lankes, M., & Bernhaupt, R. (2011). Using embodied conversational agents in video games to investigate emotional facial expressions. *Entertainment Computing*, 2, 29–37.
- Lankes, M., Bernhaupt, R., & Tscheligi, M. (2007). An experimental setting to measure contextual perception of Embodied Conversational Agents. *ACE'07 International conference on Advances in Computer Entertainment Technology*, 56–59.

- Lankes, M., Bernhaupt, R., & Tscheligi, M. (2010). Evaluating user experience factors using experiments: Expressive artificial faces embedded in contexts. In R. Bernhaupt (Ed.), *Evaluating user experience in games* (pp. 165–183). London, UK: Springer.
- Liu, C., Conn, K., Sarkar, N., & Stone, W. (2008). Physiology-based affect recognition for computer-assisted intervention of children with autism spectrum disorder. *International Journal of Human-Computer Studies*, *66*, 662–677.
- Luo, P., Kipp, M., & Neff, M. (2009). Augmenting gesture animation with motion capture data to provide full-body engagement. *IWA 2009 International Conference on Intelligent Virtual Agents*, 405–417.
- Meeeren, H., van Heijnsbergen, C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *PNAS*, *102*, 16518–16523.
- Mitchell, P., Parsons, S., & Leonard, A. (2007). Using virtual environments for teaching social understanding to 6 adolescents with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*, 589–600.
- Moore, D., Cheng, Y., McGrath, P., & Powell, J. (2005). Collaborative virtual environment technology for people with autism. *Focus on Autism and Other Developmental Disorders*, *20*, 231–243.
- Parsons, S., & Cobb, S. (2011). State-of-the-art of virtual reality technologies for children on the autism spectrum. *European Journal of Special Needs Education*, *26*, 355–366.
- Paterson, H. M., Pollick, F. E., & Jackson, E. (2002). Movement and faces in the perception of emotion. *European Conference on Visual Perception*, 118.
- Paterson, H. M., Pollick, F. E., & Sanford, A. J. (2001). The role of velocity in affect discrimination. *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, 756–761.
- Pitterman, H., & Nowicki, S. (2004). A test of the ability to identify emotion in human standing and sitting postures: The diagnostic analysis of nonverbal accuracy-2 posture test (DANVA-POS). *Genetic, Social and General Psychology Monographs*, *130*, 146–162.
- Poggi, I., Cavicchio, F., & Magno Caldognetto, E. (2007). Irony in a judicial debate: Analyzing the subtleties of irony while testing the subtleties of an annotation scheme. *Journal on Language Resources and Evaluation*, *41*, 215–232.
- Pollick, F. E., Paterson, H. M., & Mamassian, P. (2004). Combining faces and movements to recognize affect. *Journal of Vision*, *4*, 232.
- Rapcsak, S. Z., Galper, S. R., Comer, J. F., Reminger, S. L., Nielsen, L., Kaszniak, A. W., . . . Cohen, R. A. (2000). Fear recognition deficits after focal brain damage: A cautionary note. *Neurology*, *54*, 575–581.
- Ruffman, T., Sullivan, S., & Dittrich, W. (2009). Older adults' recognition of bodily and auditory expressions of emotion. *Psychology and Aging*, *24*, 614–622.
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, *115*, 102–141.
- Sakata, M., Shiba, M., Maiya, K., & Tadenuma, M. (2004). Human body as the medium in dance movement. *International Journal of Human-Computer Interaction*, *17*, 427–444.
- Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 92–120). New York, NY: Oxford University Press.
- Schouwstra, S., & Hoogstraten, J. (1995). Head position and spinal position as determinants of perceived emotional state. *Perceptual and Motor Skills*, *81*, 673–674.
- Schwartz, C., Bente, G., Gawronski, A., Schilbach, L., & Vogeley, K. (2009). Responses to nonverbal behaviour of dynamic virtual characters in high-functioning autism. *Journal of Autism and Developmental Disorders*, *40*, 100–111.
- Tartaro, A., & Cassell, J. (2006). Authorable virtual peers for autism spectrum disorders. *Workshop on Language-Enabled Educational Technology at ECAI'06 European Conference on Artificial Intelligence*.
- Tuminello, E. R., & Davidson, D. (2011). What the face and body reveal: In-group emotion effects and stereotyping of emotion in African American and European American children. *Journal of Experimental Child Psychology*, *110*, 258–274.
- Wallbott, N. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, *28*, 879–896.
- Wallraven, C., Breidt, M., Cunningham, D. W., & Bühlhoff, H. H. (2005). Psychophysical evaluation of animated facial expressions. *Applied Perception in Graphics and Visualization*, *95*, 17–24.
- Willis, M. L., Palermo, R., & Burke, D. (2011). Judging approachability on the face of it: The influence of face and body expressions on the perception of approachability. *Emotion*, *11*, 514–523.
- Yamamoto, M., & Watanabe, T. (2008). Effects of time lag of utterances to communicative actions on embodied interaction with robot and CG character. *International Journal of Human-Computer Interaction*, *24*, 87–107.

ABOUT THE AUTHORS

Stéphanie Buisine is a research scientist in Arts et Métiers ParisTech, an engineering school in Paris. She holds a master's degree in Psychology, a master's degree in Ergonomics, and a PhD in Cognitive Ergonomics from Paris Descartes University. Her research interests include human-computer interaction, innovation process, and creativity.

Matthieu Courgeon received a PhD degree in computer science from Paris-South University (2011). He now holds a PostDoc position at the European center of virtual reality. His work focuses on modeling affective and social behaviors of autonomous and interactive virtual agents. He created the MARC framework during his PhD thesis.

Aurélien Charles holds a master's degree in Industrial Design from the ENSAAMA School of Applied Arts and a master's degree in Innovation from Arts et Métiers ParisTech. He specializes in designing expressive agents, graphics, and products.

Céline Clavel received a PhD degree in Cognitive Psychology from University Paris West in 2007. Since 2010, she has been an assistant professor at Paris-South University. Her main research interest is to specify the psychology models to computer science applications, evaluate their contributions, and study their impacts on behavior.

Jean-Claude Martin is professor of Computer Science at Paris-South University and is the head of the group Cognition Perception Use (LIMSI-CNRS). He is an elected member of the executive board of the HUMAINE international association (study of emotions) and is the Editor-in-Chief of *Journal on Multimodal User Interfaces* (Springer).

Ning Tan is a user experience professional. She received her PhD degree from LIMSI-CNRS and Paris-South University for her dissertation "Posture and Space in Virtual Characters: Applications to Affective Interaction and Ambient Interaction" in 2012. She also holds a professional master's degree in Ergonomics from Paris-South University.

Ouriel Grynszpan is an associate professor in neurosciences at Université Pierre et Marie Curie, in the Emotion Center (CNRS USR 3246) at La Salpêtrière hospital in Paris. He received a PhD in computer sciences from Paris-South University in 2005. His research addresses the use of innovative technologies for cognitive disorders.