Evaluating the Effect of Emotion on Gender Recognition in Virtual Humans

Katja Zibrek* Ludovic Hoyet[†] Kerstin Ruhland[‡] Rachel McDonnell[§]

Graphics, Vision and Visualisation Group, Trinity College Dublin

Abstract

In this paper, we investigate the ability of humans to determine the gender of conversing characters, based on facial and body cues for emotion. We used a corpus of simultaneously captured facial and body motions from four male and four female actors. In our Gender Rating task, participants were asked to rate how male or female they considered the motions to be, under different emotional states. In our Emotion Recognition task, participants were asked to classify the emotions, in order to determine how accurately perceived those emotions were. We found that gender perception was affected by emotion, where certain emotions facilitated gender determination while others masked it. We also found that there was no correlation between how accurate an emotion was portrayed and how much gender information was present in that motion. Finally, we found that the model used to display the motion did not affect gender perception of motion but did alter emotion recognition.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: facial animation, emotion, gender

1 Introduction

Computer generated virtual humans are an important part of many 3D graphical applications, and are often used to invoke empathic responses from human users, due to their human form and behavior. With the advancement of real-time tracking of faces and virtual teleconferencing, producing plausible virtual characters driven by human motion is becoming increasingly important. In particular, the transmission of correctly identified and plausible emotions is highly important to avoid ambiguity or confusion. Recently, the computer graphics community has come to realize the importance of taking human perception into account when striving for plausibility and realism in animations and virtual characters. However, to date there have been few studies which have focused on the perception of computer generated facial animation driven from real humans. In this paper, we focus on the perception of gender on virtual conversing characters in an attempt to further our understanding of the communication of emotion between real and virtual humans.

In previous work [McDonnell and O'Sullivan 2010], we investigated the ability of humans to determine the sex of conversing characters, based on audio and visual cues. We found that conversational body motions carry sex information and should only be displayed with a congruent voice to avoid ambiguity. Here, we



Figure 1: Top: Female motion applied to the Woman and Man models (Happy emotion). Bottom: Male motion applied to both models (Angry emotion).

investigate further the perception of virtual conversing characters. Unlike the previous study, we are now interested in identifying if certain emotions affect gender perception (the terms *gender* and *sex* are defined in Section 4), and if emotion recognition is different on males and females. We use synchronized body and facial motion capture data in order to recreate the performance more faithfully than using either body or facial motion alone.

2 Background

Since the early work of Johansson [1973] on the perception of human motion from point-light displays (lights attached to the major joints in the body), it has been shown that humans are extremely capable of recognizing and categorizing human motion even with very little information. Several studies focused on understanding how gender is expressed through human motion. For instance, it has been shown that gender can be perceived through body motion alone (e.g. [Kozlowski and Cutting 1977]), with hip sway often indicating female motion and shoulder movement indicating male motion [Johnson and Tassinary 2005]. Other studies also focused on the recognition of gender from facial motion, using either androgynous faces [Hill and Johnston 2001; Morrison et al. 2007] or comparisons with point-light displays [Hill et al. 2003]. These studies found that gender can be recognized from facial motion only, but that specific cues, such as excessive nodding, blinking and overall amount of movement, play an important role in recognizing the female gender [Morrison et al. 2007]. Similarly, Hill et al. [2001] showed that head motion alone is less useful for discriminating sex than facial movements.

As humans have a high sensitivity to emotions conveyed by other humans, researchers also investigate how emotions are perceived

^{*}e-mail:katjazibrek@tcd.ie

[†]e-mail:hoyetl@scss.tcd.ie

[‡]e-mail:ruhlandk@tcd.ie

[§]e-mail:Rachel.McDonnell@scss.tcd.ie

and recognized in human facial expressions, body poses and motions. The studies usually investigate six basic emotions: Anger, Happiness, Sadness, Fear, Surprise and Disgust; according to Ekman's classification [1992]. Studies using point-light displays for recognition of movement show that emotions can be identified from either body motion [Atkinson et al. 2004] or facial motion [Bassili 1978]. Additionally, recognition rates are improved when body and facial motions are combined [Clavel et al. 2009], showing that both facial and body cues have a specific role in conveying emotion. Research also shows that there are particular areas of the face [Schyns et al. 2009] and certain body motions [De Meijer 1989] that convey most of the information about the portrayed emotions, e.g. corners of the mouth are important in recognizing happiness, which suggests that the recognition of emotion can be improved when this behavioral data is present.

Gender stereotypes impact both the production and perception of emotions [Brewer 1988; Devine 1989]. For instance, Western stereotypes usually describe women as more affiliative, and more likely to show happiness, and men as being more dominant, and therefore more likely to show anger [Brody and Hall 2000; Fischer 1993; Lewis 2000]. Studies conducted on static images [Hess et al. 2004] also found differences in the way the same emotional expressions are perceived on female and male static faces. This is particularly the case for happiness and anger, where women were rated as more angry than men, and men as happier than women. Hess et al. [2004] attributed this finding to the contrast effect, since a woman showing anger or a man expressing happiness might violate the viewers expectation of gender stereotype and therefore the incongruence between gender and emotion would intensify the perceived emotional expression. Using video references of female and male actors portraying emotions, Battocchi et al. [2005] found that emotions are better recognized on female actors overall, especially anger and sadness. While both studies explored the interaction between gender and emotion, neither separated appearance of the model from dynamic motion. They used either video references, where the sex of the actor was apparent, or static images and thus left out the component of dynamic motion. The study of Johnson et al. [2011] combined both aspects by exploring gender recognition bias on the perception of throwing a ball under different emotions using point-light displays. They found that an angry throw is perceived as more male and a sad throw as more female, which confirms previous findings.

In contrast to previous work, the main aim of this paper is to explore gender recognition for conversational emotions using both body and facial motions on realistic virtual characters. According to the above mentioned research, gender can be recognized from either body or facial motion, but none of these studies explored how gender judgments differ for different emotional expressions using motion cues for both the face and the body. By using a Man and a Woman model to display the motions of both male and female actors, we explore how gender recognition is affected by the visual appearance of the character, especially since previous work has shown that the choice of the model affects the perception of motion [Hodgins et al. 1998; Chaminade et al. 2007].

3 Stimuli Creation

To address the question of gender and emotion perception on virtual characters, we chose to use natural motions captured on real actors, as they capture the subtleties and timings of real movements. The motion capture data used in this paper was acquired using a 21 camera Vicon optical system, where 52 markers were placed on the body and 36 markers on the face. We did not capture eye or finger motion. The motion capture data included 4 female and 4 male trained actors, acting out short sentences portraying four of the ba-

sic emotions (anger, fear, happiness and sadness) as well as neutral sentences. We selected the four emotions evaluated in [Ben-David et al. 2011] where they define a validated list of affective sentences for spoken emotion identification e.g. "Get out of my room!" for anger or "It's a beautiful day outside!" for happiness. From this validated list, we chose three sentences for each emotion (including Neutral). Since we wanted to examine the effect of motion alone, the voices of the actors, although captured, were not used in these experiments.

To evaluate how the character's model influences the perception of gender, two virtual characters of different sex were used to display the motions of our actors. We used congruent (where character sex matched actor sex) and in-congruent (where character sex did not match actor sex) stimuli (Figure 1). The captured body markers were used to compute the joint angle animations and mapped onto the virtual characters in Autodesk 3ds Max 2012. The facial motion was directly exported as 3D marker motion and stabilized by removing the movement of the head. A bone-based approach that used linear blend skinning was used to drive the facial geometry. In order to retarget the motion from actor to model, the markers were aligned to the head and then automatically adjusted to the position of the bones of the face of the character. The characters' facial bones were then constrained to their corresponding optical markers to produce the animation.

The character was displayed in the center of the screen, facing forward at the beginning of each clip. We selected a medium close up view as we wanted to provide information from both the body and facial movements of the actors (Figure 1). Movies were presented at 30 frames per second and at 1240×900 resolution on a 24-inch LCD screen. Two hundred and forty 3-second clips were generated in total: 2 Models (Woman, Man) \times 8 Actors (4F, 4M) \times 5 Acted Emotions (Anger, Fear, Happy, Sad + Neutral) \times 3 Sentences.

4 Experiment

The experiment was divided into two blocks. In the first block, participants rated the gender of the character based on their perception of motion. In the second block, they were asked about the type of emotion the virtual characters were expressing. In both blocks, participants were presented with the same set of 240 video clips depicting virtual characters conveying different emotional sentences. Trials lasted 3 seconds each and participants were not presented with audio.

Gender Rating. In the first block, participants were asked to rate the gender of the character, based on motion cues. They were informed that the displayed motion would not necessarily match the sex of the model. After each clip, they rated Gender on a 5-point Likert scale from 1 (Very Male) to 5 (Very Female) using the number keys on a keyboard. Half of the participants viewed the Man model stimuli first and the other half viewed the Woman model first. Clips were randomized to avoid ordering effects.

We decided to use the term *gender* for the rating task and the term *sex* for the actual underlying motion (male, female) based on the distinction between *gender* as a sociocultural construct associated with maleness and femaleness and *sex* as a demographic category [Unger 1979]. Because we instructed the participants to pass judgment on a 5-point scale of how female or male the motion was, and did not ask to categorize the motion as either male or female, our task is closer to the sociocultural use of the term. However, our study does not presuppose that the origin of the distinction in motion between males and females is of sociocultural or biological nature, as pointed out by Deaux [1993].

Emotion Recognition. In the second block, participants categorized and then rated the intensity of the displayed emotion. They were asked "Which of the 5 listed emotions is the character expressing?" and selected an emotion by pressing the corresponding key on the keyboard, marked with A (Anger), F (Fear), H (Happy), N (Neutral) and S (Sad). We instructed them to use the Neutral condition only when they believed that the character did not express any emotion. They were also informed at the beginning of the experiment that one of the listed emotions would always be portrayed in each clip. Immediately following their response, participants were then asked to "Rate the intensity of motion", except when selecting Neutral for which an intensity scale was not appropriate. They rated the intensity on a 5-point Likert scale from 1 (not intense at all) to 5 (extremely intense) using the number keys on a keyboard.

Sixteen participants (8M-8F; average age: 31 ± 6.6) took part in each block of this experiment. Fifteen participants completed both blocks (7M-8F) and two additional participants completed just one block each. They were recruited mainly from the university, had normal or corrected to normal vision, and were all naïve to the purpose of the experiment.

5 Results

Results for both the Gender Rating and Emotion Recognition tasks were analyzed using repeated measures ANalysis Of VAriance (ANOVA) with within-subjects factors: *actor sex, model*, and *acted emotion*, and between-groups factor *participant sex* on all of our data. In all cases, we averaged over the data for the four actors for actor sex, and over the three sentences. No effect of participant sex was observed for any tests conducted so this is not discussed further. Post hoc analysis was conducted with Newman-Keuls tests for comparison of means.

5.1 Gender Rating

We first investigated the results for gender ratings, and found that participants were able to distinguish between male (Avg: 2.53 ± 0.03) and female (Avg: 3.49 ± 0.05) motions overall (F(1,14) = 246.64, $p\approx0$). This ability to rate the gender of the motion was unaffected by the character model used to present the motion. However, ratings of the gender of the motion were affected by the acted emotion (F(4, 56)=17.031, $p\approx0$) (Figure 2). In general, Anger (Avg: 2.47 ± 0.11) was considered significantly more male than Fear (Avg: 3.09 ± 0.05), Happy (Avg: 3.24 ± 0.05), Sad (Avg: 3.2 ± 0.07), or Neutral (Avg: 3.03 ± 0.08).

An interaction also occurred between actor sex and emotion (F(4, 56)=54.215, $p\approx 0$), where post-hoc analysis showed that all female emotions were rated as significantly more female than their male counterparts (p<0.0003 in all cases). For females, Happy was rated significantly more female than all other emotions (p<0.0002 in all cases). Fear and Sad were rated equally as the next most female (p<0.0008 in all cases), followed by Neutral (p<0.0008 in all cases), and finally Anger which was rated as significantly less female than all others (p<0.0003 in all cases). For males, Sad and Neutral were rated equally as the most female of all emotions (p<0.009 in all cases). Fear was considered the next most female (p<0.009 in all cases), followed by Happy (p<0.003 in all cases), and finally Anger was rated as the least female of all male emotions (p<0.003 in all cases).

To investigate the meaning of these results further, we conducted one-sample *t*-tests to determine if participants could recognize actor sex for each emotion (comparison with the constant 3 which related to ambiguous gender). We found that for Anger, participants could not determine the gender of the female motions, but could for the

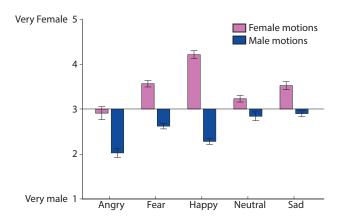


Figure 2: Interaction between Actor Sex and Acted Emotion on Gender ratings.

male motions (t(15) = -9.63, $p \approx 0$). For Neutral and Sad, male motions did not convey gender information, but female motions did (Sad: t(15) = 5.93, $p \approx 0$; Neutral: t(15) = 3.14, p < 0.07). For Happy and Fear, male motions were recognized as male (t(15) = -10.724, $p \approx 0$) and female motions as female (t(15) = 13.809, $p \approx 0$).

Finally, the model used to present the emotions did not affect ratings (no interaction between model and emotion, or three way interactions with model). These results imply that appearance does not affect gender perception of motion, and that gender cues are different for different emotions.

5.2 Emotion Recognition

We also investigated the results for emotion recognition, and found that overall, participants were very accurate in identifying the acted emotions (68% accuracy on average), with 20% being the probability of recognition being due to chance. However, we found a main effect of emotion (F(4, 60)=36.668, $p\approx 0$) and post-hoc analysis showed that Anger and Happy were significantly more recognized than any other emotion (p<0.0002 in all cases). Sad and Fear were rated equally accurately, and Neutral was the least recognized emotion (p<0.009 in all cases). Confusion matrices show where misclassifications occurred (Table 1).

There was also a main effect of actor sex (F(1,15) = 192.57, $p \approx 0$), which showed that participants were more accurate at emotion recognition when viewing female (79.48%±1.36) than male (67.09%±1.48) motion. This is consistent with previous work which showed that females are more accurate at conveying emotion than males [Battocchi et al. 2005]. An interaction between actor sex and acted emotion (F(4, 60)=30.473, $p \approx 0$) showed that this was due to fact that all emotions except Anger and Happy were more easily recognized on the female actor (p < 0.04 in all cases).

We also found that the model used for viewing the motion had an effect on accuracy (F(1,15) = 9.074, p < 0.009), with motion recognition being more accurate on the Man (74.74 % \pm 1.45) than on the Woman model (71.83% \pm 1.41).

We also found an interaction between acted emotion and model, which was due to the fact that Fear and Sad were detected more easily on the Man than on the Woman (p < 0.0002 for both). This result implies that appearance affects emotion perception, and care should be taken when choosing a virtual model to convey emotion.

		Female Model					Male Model				
Actor Sex	Emotion	Angry	Fear	Нарру	Neutral	Sad	Anger	Fear	Нарру	Neutral	Sad
Female Actor	Angry	97.9%	0.5%	0%	0.5%	1.1%	93.2%	3.7%	0%	0.5%	2.6%
	Fear	9.4%	73.4%	2.1%	1.6%	13.5%	3.6%	84.9%	0.5%	1.6%	9.4%
	Нарру	6.3%	1.0%	90.1%	0.5%	2.1%	4.7%	0.5%	88.0%	3.1%	3.7%
	Neutral	6.8%	2.6%	4.7%	62.0%	23.9%	3.6%	4.2%	1.6%	56.2%	34.4%
	Sad	7.3%	7.3%	1.6%	15.6%	68.2%	1.6%	9.4%	0.5%	7.8%	80.7%
Male Actor	Angry	97.4%	1.6%	0.0%	0.5%	0.5%	94.8%	4.2%	0.5%	0.0%	0.5%
	Fear	28.1%	46.4%	3.6%	10.4%	11.5%	20.3%	58.3%	3.1%	7.3%	11.0%
	Нарру	10.4%	0.5%	86.0%	2.6%	0.5%	7.8%	0.0%	87.5%	3.7%	1.0%
	Neutral	9.9%	1.6%	10.9%	47.9%	29.7%	8.3%	2.6%	7.8%	44.8%	36.5%
	Sad	12.0%	6.8%	6.8%	25.5%	48.9%	5.7%	11.5%	1.6%	22.4%	58.8%

Table 1: Confusion matrix for participants emotion selection for each condition in our experiment. The emotion selected significantly most often is highlighted.

5.3 Emotion Intensity

Results for emotion intensity were also tested, where we included data only from correctly identified emotions. We found a main effect of emotion (F(3, 42)=38.202, $p\approx 0$), where post-hoc analysis showed that Anger, Sad, Happy and Fear were rated as equally intense, while Sad was rated as significantly less intense, as expected (p<0.0002 in all cases). We also found that actor's sex had an effect on the perceived intensity of emotions (F(1,14)=9.6061, p<0.008) where male actors were rated as more intense than females overall. Finally, we found an interaction between actor sex and emotion (F(3,42)=19.019, $p\approx 0$), where male actors were considered more intense than females for Anger (p<0.0002) and less for Sad (p<0.04) emotions (Figure 3).

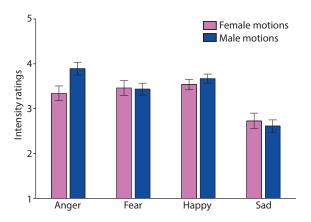


Figure 3: Interaction between Actor Sex and Acted Emotion on Intensity ratings.

5.4 Correlation

Since some emotions were easier to recognize than others we were interested to determine whether recognition accuracy was related to the ability to recognize the underlying actor's gender. We therefore tested for correlations between emotion and gender recognition using Pearson's coefficient across averages over emotions. We found no significant correlation for any emotion besides a low correlation for Happy $(r=0.17;p\approx0)$. This implies that for most emotions gender recognition is not necessarily facilitated by how accurate an emotion is portrayed.

6 Discussion

We have conducted an experiment to determine the effect of emotion on the perception of gender for virtual conversing characters. Our main finding is that gender information was nullified but not overridden by certain stereotypical emotions. For example, when female actors expressed a stereotypically masculine emotion (Anger), the resulting animation was not perceived as male, rather, it was considered to be ambiguous in gender. Similarly, when the Sad emotion was expressed by male actors, their motion was perceived as ambiguous. Surprisingly, Fear and Happy emotions accentuated gender information and were perceived as highly female when acted by females, and highly male when acted by males.

We also tested the effect of the appearance of the character and found that surprisingly, appearance did not accentuate gender in congruent pairs (e.g. male motion applied to the Man model), nor did it alter gender perception of in-congruent pairs. However, we did find that participant accuracy of emotion detection was affected by the model used to convey it. In particular, Sad and Fear emotions were detected more easily on the Man model than on the Woman. This result could be due to the fact that the model's face affects the perception of its emotion. For example, if a model's eyebrows are lowered in its neutral state, it may appear sad (see behavioral data for different emotional expressions in [Schyns et al. 2009]), and all subsequent motions that are applied to that model will be affected. Therefore, care should be taken when choosing a virtual model to display emotion to ensure that the model's face itself does not carry any specific emotional information. On the other hand, the result could be due to motion mapping, skinning, or retargetting effects causing differences between the models. Therefore, in future work we will assess the effect of emotion perception on a range of models to determine if the result is due to gender differences or specifics of the models used.

In our experiments, participants were asked to rate the motion only and ignore the appearance. However, we were interested to determine if the model inadvertently alters the perception of gender of the motion that it is animated with. We found this not to be the case and therefore suggest that in-congruent pairs should be avoided where possible, to avoid ambiguity or unnatural results.

One limitation of our study is that we did not control for the intensity of emotion in the stimuli creation. This was due to the fact that we were interested in natural expression of emotion and felt that controlling for intensity would result in unnatural expressions (e.g., Anger expressed at a low intensity, or Sad expressed at a high intensity). However, we did ask participants to report how intense they found the emotions to be, in order to explore the possibility that higher intensity results in stronger gender cues. Our results in-

dicate that gender perception is emotion specific and not due to how intense or accurately perceived the emotion is. Further studies will investigate the correlation between emotion recognition, intensity and gender perception.

Acknowledgements

This work was sponsored by Science Foundation Ireland as part of the Cartoon Motion and Captavatar projects.

References

- ATKINSON, A. P., DITTRICH, W. H., GEMMELL, A. J., YOUNG, A. W., ET AL. 2004. Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception* 33, 6, 717–746.
- BASSILI, J. N. 1978. Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology Human Perception and Performance* 4, 373–379.
- BATTOCCHI, A., PIANESI, F., AND GOREN-BAR, D. 2005. A first evaluation study of a database of kinetic facial expressions (dafex). In *Proceedings of the 7th International Conference on Multimodal Interfaces*, ICMI '05, 214–221.
- BEN-DAVID, B. M., VAN LIESHOUT, P. H. H. M., AND LESZCZ, T. 2011. A resource of validated affective and neutral sentences to assess identification of emotion in spoken language after a brain injury. *Brain Injury* 25, 2, 206–220.
- Brewer, M. B. 1988. A dual process model of impression formation. Lawrence Erlbaum Associates, Inc.
- Brody, L. R., AND HALL, J. A. 2000. Gender, emotion, and expression. In *Handbook of Emotions*, Guilford, M. Lewis and J. Haviland-Jones, Eds., 338–349.
- CHAMINADE, T., HODGINS, J., AND KAWATO, M. 2007. Anthropomorphism influences perception of computer-animated characters' actions. *Social Cognitive and Affective Neuroscience* 2, 3, 206–216.
- CLAVEL, C., PLESSIER, J., MARTIN, J.-C., ACH, L., AND MOREL, B. 2009. Combining facial and postural expressions of emotions in a virtual character. In *Proc. of the 9th International Conference on Intelligent Virtual Agents*, IVA '09, 287–300.
- DE MEIJER, M. 1989. The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal behavior 13*, 4, 247–268.
- DEAUX, K. 1993. Commentary: Sorry, wrong number: A reply to gentile's call. *Psychological Science* 4, 2, pp. 125–126.
- DEVINE, P. G. 1989. Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology* 56, 1, 5–18.
- EKMAN, P. 1992. An argument for basic emotions. *Cognition & Emotion 6*, 3-4, 169–200.
- FISCHER, A. H. 1993. Sex differences in emotionality: Fact or stereotype? *Feminism & Psychology 3*, 3, 303–318.
- HESS, U., ADAMS, R. B., AND KLECK, R. E. 2004. Facial appearance, gender, and emotion expression. *Emotion 4*, 4, 378–388.

- HILL, H., AND JOHNSTON, A. 2001. Categorizing sex and identity from the biological motion of faces. *Current Biology* 11, 11, 880–885.
- HILL, H., JINNO, Y., AND JOHNSTON, A. 2003. Comparing solid-body with point-light animations. *Perception* 32, 5, 561–566.
- HODGINS, J. K., O'BRIEN, J. F., AND TUMBLIN, J. 1998. Perception of human motion with different geometric models. *IEEE Trans. on Visualization and Computer Graphics* 4, 4, 307–316.
- JOHANSSON, G. 1973. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics 14*, 2, 201–211.
- JOHNSON, K. L., AND TASSINARY, L. G. 2005. Perceiving sex directly and indirectly meaning in motion and morphology. *Psy*chological Science 16, 11, 890–897.
- JOHNSON, K. L., MCKAY, L. S., AND POLLICK, F. E. 2011. He throws like a girl (but only when he's sad): Emotion affects sexdecoding of biological motion displays. *Cognition* 119, 2, 265 – 280.
- KOZLOWSKI, L. T., AND CUTTING, J. E. 1977. Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics* 21, 6, 575–580.
- LEWIS, K. M. 2000. When leaders display emotion: how followers respond to negative emotional expression of male and female leaders. *Journal of Organizational Behavior 21*, 2, 221–234.
- MCDONNELL, R., AND O'SULLIVAN, C. 2010. Movements and voices affect perceived sex of virtual conversers. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, 125–128.
- MORRISON, E. R., GRALEWSKI, L., CAMPBELL, N., AND PENTON-VOAK, I. S. 2007. Facial movement varies by sex and is related to attractiveness. *Evolution and Human Behavior* 28, 3, 186–192.
- SCHYNS, P. G., PETRO, L. S., AND SMITH, M. L. 2009. Transmission of facial expressions of emotion co-evolved with their efficient decoding in the brain: behavioral and brain evidence. *PLoS One* 4, 5, e5625.
- UNGER, R. K. 1979. Toward a redefinition of sex and gender. *American Psychologist 34*, 11, 1085–94.