

Body Expressions Influence Recognition of Emotions in the Face and Voice

Jan Van den Stock and Ruthger Righart
Tilburg University

Beatrice de Gelder
Tilburg University, Massachusetts General Hospital, and
Harvard Medical School

The most familiar emotional signals consist of faces, voices, and whole-body expressions, but so far research on emotions expressed by the whole body is sparse. The authors investigated recognition of whole-body expressions of emotion in three experiments. In the first experiment, participants performed a body expression-matching task. Results indicate good recognition of all emotions, with fear being the hardest to recognize. In the second experiment, two alternative forced choice categorizations of the facial expression of a compound face-body stimulus were strongly influenced by the bodily expression. This effect was a function of the ambiguity of the facial expression. In the third experiment, recognition of emotional tone of voice was similarly influenced by task irrelevant emotional body expressions. Taken together, the findings illustrate the importance of emotional whole-body expressions in communication either when viewed on their own or, as is often the case in realistic circumstances, in combination with facial expressions and emotional voices.

Keywords: emotion perception, body expression, face expression, voice prosody

Imagine yourself taking a walk in the woods with a friend. You're talking to each other and suddenly he sees a snake approaching. His behavior changes immediately: his eyes open widely, his shoulders move backward, and his tone of voice changes. Over the last decade considerable progress has been made in understanding the functional and neuro-anatomical basis of human emotions. The combined findings from psychophysical investigations, lesion studies, and brain imaging in neurologically intact observers have already provided a wealth of insights in how viewers process emotional information. Yet, it is noteworthy that our present understanding of how emotions are processed in humans is almost entirely based on studies investigating the perception of facial expressions (Adolphs, 2002).

Considering the emotional value of bodily expressions, it is somewhat surprising that the study of perception of whole-body expressions lags so far behind that of facial expressions. Whole-body expressions provide information about the emotional state of the producer, but also signal his action intentions. For example, a fearful body expression can signal the presence of a threat, but also

how the producer intends to deal with it: flee, fight, or freeze. Therefore, body expressions reveal a close link between emotion and (adaptive) behavior. Despite the early work of Darwin (1872), who described in detail the body expressions of many different emotions, there have been only a few isolated studies on human body postures in the past decades (Argyle, 1988; Ekman, 1965; Sprengelmeyer et al., 1999).

Our ability to perceive these emotional behaviors and how they are represented in the brain are now becoming important research topics. The findings so far have revealed striking similarities between how we process facial and whole-body emotions (de Gelder, 2006).

For instance, at the behavioral level, some of the well-explored perceptual mechanisms involved in face processing also play a role in perception of bodies. Faces and bodies seem to be processed as invariant configurations inducing so called configural processing strategies, whereas other complex stimuli are more processed as an assemblage of features. Configural processing is often measured by the inversion effect (configural stimuli presented upside-down are more difficult to recognize than other complex inverted stimuli), and this effect has recently also been reported for bodies (Reed, Stone, Bozova, & Tanaka, 2003). Similar to the face inversion effect, the body inversion effect has also been measured with event related potentials (ERP). A strong ERP inversion effect similar to that obtained for faces was observed for bodies (Stekelenburg & de Gelder, 2004).

Recent developmental findings now also underscore the important role of perceiving bodies for infants. A preferential processing of either faces or bodies might be a function of the distance from the stimulus. This suggests that if a face is present at close range, especially the eyes are important, but when the distance increases, the configural properties of the whole face play a role (Johnson, 2005). This argument can be extended to whole bodies and suggests that whole-body expressions are preferentially processed

Jan Van den Stock and Ruthger Righart, Cognitive and Affective Neuroscience Laboratory, Tilburg University, the Netherlands; Beatrice de Gelder, Cognitive and Affective Neuroscience Laboratory, Tilburg University, the Netherlands, and Martinos Center for Biomedical Imaging, Massachusetts General Hospital, and Harvard Medical School, Charlestown, Massachusetts.

Research was funded by a HFSP (Human Frontiers Science Program), an NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek), and a European Commission (COBOL) grant. We thank all participants for their cooperation.

Correspondence concerning this article should be addressed to Beatrice de Gelder, Cognitive and Affective Neuroscience Laboratory, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, the Netherlands. E-mail: b.degelder@uvt.nl

when the perceiver is further away from the stimulus. In line with this, behavioral data indicate the existence of specific expectations about the canonical properties of static faces and bodies at around 18 months (Slaughter, Stone, & Reed, 2004) and of dynamic bodies at 3 months (Bertenthal, Proffitt, & Kramer, 1987). Furthermore, ERP recordings provide evidence for similar processing of the configuration of faces and bodies at 3 months of age (Gliga & Dehaene-Lambertz, 2005).

At the functional neuro-anatomical level, a brain area in lateral occipital cortex has been described as responding selectively to neutral bodies or body parts (Downing, Jiang, Shuman, & Kanwisher, 2001). The fact that this area is very close to the motion sensitive area MT may explain its sensitivity to movement (Astafiev, Stanley, Shulman, & Corbetta, 2004, but see Peelen & Downing, 2005; Schwarzlose et al., 2005). Also, recent observations indicate significant proximity between faces and bodies in fusiform cortex (Schwarzlose, Baker, & Kanwisher, 2005) consistent with the finding that fearful bodies activate the face area in middle fusiform cortex (de Gelder et al., 2004; Hadjikhani & de Gelder, 2003) and the finding that watching video images of angry hands and angry faces activate largely overlapping brain areas (Grosbras & Paus, 2006).

Recently, we proposed a model for the underlying circuitry of perception of emotional body language (de Gelder, 2006), advocating a two-system network, with the amygdala at the core: a primary subcortical one (including amygdala, striatum, pulvinar, and superior colliculus) involved in rapid automated reflex-like perception of whole-body expressions and a more cortically based one involved in explicit recognition. The latter system comprises the amygdala, superior temporal sulcus, presupplementary motor area, inferior parietal lobule, and inferior frontal gyrus (Grèzes, Pichon, & de Gelder, 2007; Grosbras & Paus, 2006).

The present study investigates emotional body postures, how they are perceived, and what their influence is on the recognition of facial and vocal expressions of emotion. In Experiment 1, we investigated how well emotions are recognized from bodily expressions. In Experiment 2, we addressed the issue of synergies between facial expressions and bodily expressions. In Experiment 3, we explored the impact of bodily expressions on recognition of emotional voices.

Experiment 1: Recognition of Bodily Expressions

The goal of this experiment was to test recognition of body expressions with a newly developed set of emotional body images. We asked participants to match a validated set of whole-body expressions in a two-alternative forced choice task. We used a matching task instead of a naming or categorization task because we wanted to investigate how well the different emotions are recognized on the basis of similarities with other stimuli in the same category and not mediated by the use of verbal labels.

Method

Participants. A total of 17 neurologically intact volunteers, between the age of 18 and 28 years (mean age = 21.3 years), participated in the experiment.

Materials and procedure. Materials consisted of 72 gray-scale photographs representing semiprofessional actors (half male) ex-

pressing different emotions with their whole body (anger, fear, happiness, and sadness) but with the face blurred. Selection of materials for use in the present experiment was based on the results of a pilot study in which the images were presented one by one on a screen and shown for 4000 ms with a 4000-ms interval. Participants were instructed to categorize each stimulus in a forced-choice procedure choosing one among four emotion names as quickly and as accurately as possible and indicating the response on an answering sheet. For use in the present study, we only used images recognized above 70% accuracy.

A stimulus consisted of a target picture presented at the top and two probes left and right underneath (see Figure 1A for an example). There were always three different identities; all three of the same gender and one of the probes had the same expression as the target. We balanced the design in such a way that, for example, when fear was the target expression, there were two trials (one with male actors and one with female actors) with an angry distracter, two trials with a happy distracter, and two trials with a sad distracter. A total of 72 images was used, arranged in 24 trials (4 emotion categories \times 3 distracter categories \times 2 genders). To avoid identity-based matching, we used three different identities on each trial.

Stimuli were presented on a computer screen, and participants were requested to match (as accurately and fast as possible) one of the bottom pictures to the one on top, based on similarity of expressed emotion. No instructions were given about which emotions could be expected on each particular trial. They responded by pressing the corresponding button, indicating their choice for the left or right probe. The stimulus was presented until response. During the 1000-ms intertrial interval, a blank screen was shown.

Results

Mean accuracy data are shown in Figure 1B, mean reaction time data in Figure 1C. One-sample *t* tests show that recognition of all body emotions is above chance level (50%), $t(16) \geq 12.33$, $p \leq .001$.

A repeated measures analysis of variance (ANOVA) was carried out with expression (4 levels: anger, fear, happiness, and sadness) as a within-subjects variable. This resulted in a significant effect, $F(3, 48) = 10.37$, $p < .001$. Bonferroni corrected post hoc paired samples *t* tests showed significant differences between anger and sadness, $t(16) = 3.79$, $p < 0.002$; fear and happiness, $t(16) = 4.40$, $p < 0.001$; and fear and sadness, $t(16) = 5.22$, $p < 0.001$.

A repeated measures ANOVA on the same data, but with expression of the distracter as a within-subjects factor also showed a significant effect, $F(3, 48) = 4.69$, $p < .006$. Bonferroni corrected post hoc paired samples *t* tests showed significant differences between anger and sadness as distracters, $t(16) = 3.85$, $p < 0.001$; and fear and sadness as distracters, $t(16) = 3.05$, $p < 0.008$. This is in line with the findings of the analysis with target body expression as a within-subjects variable, because it shows that angry and fearful bodily expressions are recognized less accurately than sad bodily expressions.

In order to find out which emotion the expression fear was most often confused with, we calculated the number of errors as a function of distracter emotion on the trials where fear was the target emotion. Seventy-nine percent of the errors were made when

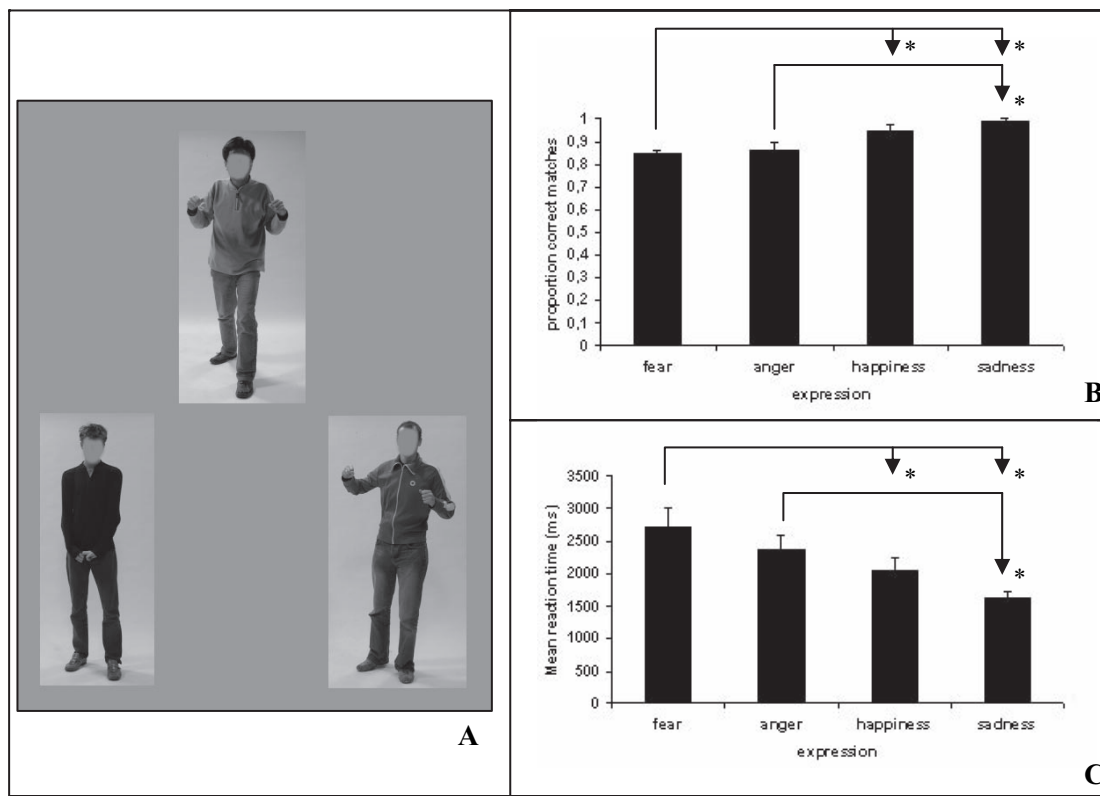


Figure 1. (A) Stimulus example of a trial in Experiment 1, showing an angry expression on top (target) and bottom right and a sad expression on the bottom left. Graphs show accuracy (B) and mean reaction times (C) as a function of expression. Error bars represent 1 SEM. * $p < .005$.

anger was the distracter, indicating fear was most frequently confused with anger.

We calculated the median reaction times per participant per condition and conducted a repeated measures ANOVA with expression (4 levels) as a within-subjects variable. This showed a significant effect, $F(3, 48) = 17.18, p < .001$. Bonferroni corrected post hoc paired sampled t tests revealed significant differences between anger and sadness, $t(16) = 5.10, p < 0.001$; fear and happiness, $t(16) = 3.92, p < 0.001$; and fear and sadness, $t(16) = 3.94, p < 0.001$.

Discussion

The results of Experiment 1 indicate that the stimuli from our newly developed set body expressions are well recognizable without the help of verbal labels. The data also provide evidence for fear as the most difficult bodily expression to recognize in a forced choice paradigm. This finding has also been reported for facial expressions (Milders, Crawford, Lamb, & Simpson, 2003). Fearful expressions can be variable, depending on the kind of threat: one can be afraid of the dark, of getting hit, of making a public appearance, of being rejected, and so forth. These different kinds of fear are associated with different defensive behaviors. This may explain why fearful whole-body expressions are more difficult to recognize. However, Atkinson, Dittrich, Gemmell, and Young (2004) presented static and dynamic whole-body expressions

(face-blurred) at three levels of intensity and both in full-light and point-light displays. They asked participants to verbally label the stimuli in a five-alternative forced choice task (anger, disgust, fear, happiness, and sadness). For the static full-light displays, they found anger to be more poorly recognized than fear, happiness, and sadness, with little difference between the latter. In the present study, we found no significant difference between angry and fearful bodies, but fear was more poorly recognized than happiness and sadness (as indicated by both accuracy and reaction time data). Apart from the methodological differences (like, for example, the number of presented emotions and the type of task), the differences between emotions reported by Atkinson and colleagues may reflect differences in how well the stimuli are recognized. We accounted for this possibility by selecting the photographs that were equally well recognized in a pilot study.

Experiment 2: The Influence of Bodily Expressions on Recognition of Facial Expressions

Only one study has investigated the combined perception of human facial and bodily expressions (Meeren, van Heijnsbergen, & de Gelder, 2005). Participants were presented compound images of faces on bodies and their emotional content was either congruent or incongruent. The participants' task was to categorize the facial expression. Electrical brain responses were measured with an electroencephalogram. The behavioral results showed that re-

sponses were more accurate and faster when face and body expressed the same emotion. The ERP data provided evidence for an early perceptual integration of emotions expressed by face and body (around 115 ms post stimulus onset). We examine whether the effects observed by Meeren et al. (2005) can be replicated with fear and happiness, instead of fear and anger, as emotions. Furthermore, morphed faces were used in order to test whether individuals use information from bodies differently when facial expressions are ambiguously positioned between fear and happy.

Method

Participants. Participants were 14 first-year psychology students (mean age = 19.1 years).

Materials and procedure. Gray-scale photographs of a male actor with a fearful and happy body expression were selected from our own validated database (recognized correctly 100% and 90%,

respectively). One identity of the Ekman and Friesen (1976) facial expressions database was selected. We used the happy and fearful expression as extremes to create a 5-step continuum between the two expressions. The morphing of the expressions was done according to the procedure developed by Benson and Perrett (1991). The faces were edited in size and pasted on the body to create a realistically looking “identity” (see Figure 2B, for examples). Every facial expression was paired with every bodily expression. This resulted in 10 compound stimuli: the five facial expressions pasted on the happy bodily expression and the same five faces pasted on the fearful bodily expression.

All compound stimuli were presented on a computer screen 15 times in random order in three identical blocks. Presentation time was 150 ms, after which a blank screen appeared. Participants were instructed to indicate whether the face expressed fear or happiness. Intertrial interval was 2000 ms.

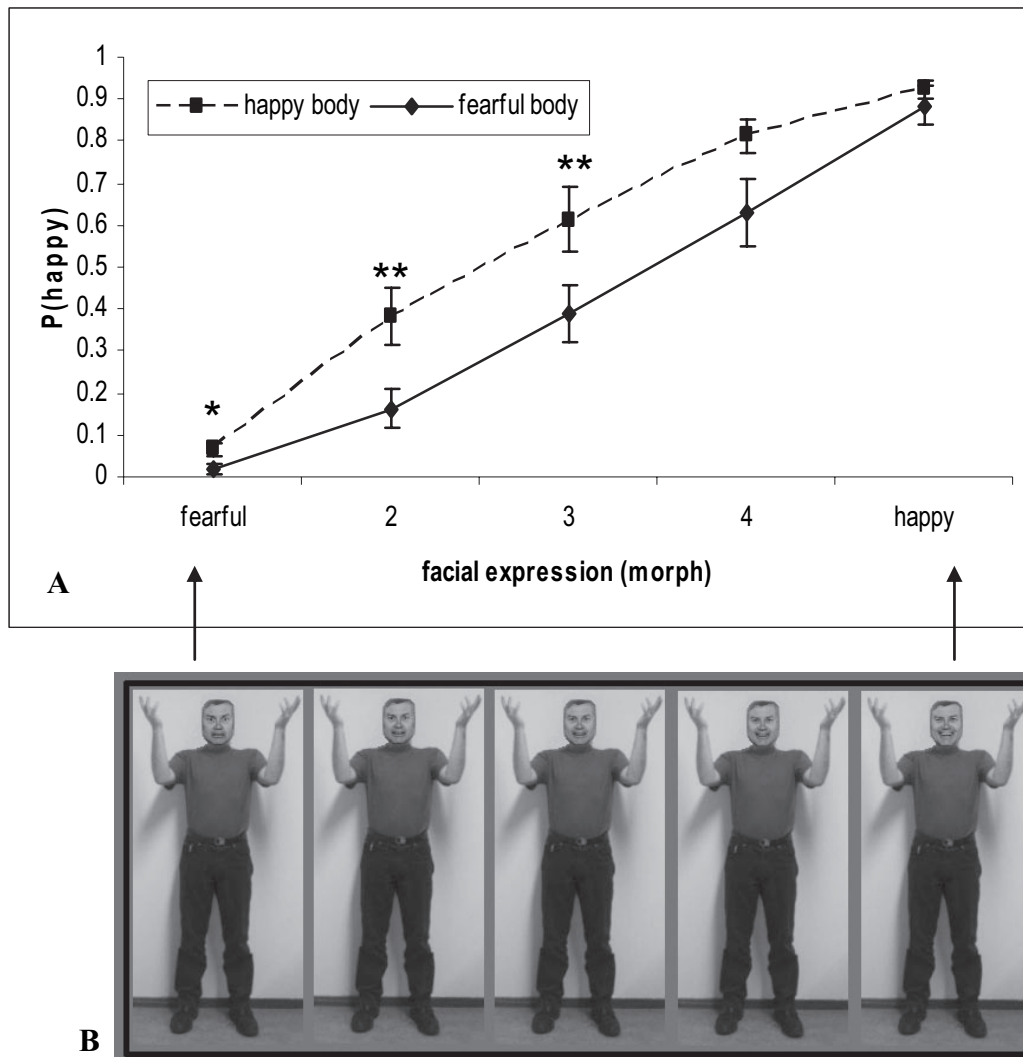


Figure 2. (A) Mean proportion happy responses in Experiment 2 as a function of facial expression. Error bars represent 1 SEM around the mean. * $p < .01$; ** $p < .001$. (B) Stimulus examples of Experiment 2, showing a happy body expression with a morphed face (left: 100% fearful; right: 100% happy).

Results

The proportion of *happy* responses was calculated for each participant and for each compound stimulus. Results are displayed in Figure 2A. A 5×2 repeated measures ANOVA was carried out with face (5 levels) and body (2 levels) as within-subjects variables. This revealed a main effect of face, $F(4, 52) = 106.65, p < .001$, body, $F(1, 13) = 37.56, p < .001$ and a significant interaction, $F(4, 52) = 4.78, p < .002$. To follow up on the interaction effect we compared for each of the 5 facial expressions, the difference between the proportions “happy” responses as a function of the accompanying bodily expression. Bonferroni corrected *t* tests showed a significant difference on three adjacent levels of the facial expression continuum, starting from the fear end ($p < .006, p < .001$, and $p < .001$, respectively). The other two differences were not significant ($p < .018$ and $p < .265$).

A trend analysis showed there was a linear trend in the face factor, $F(1, 13) = 554.33, p < .001$, indicating the distances between the face morphs were perceived as equal. The body \times face interaction showed a quadratic trend, $F(1, 13) = 23.65, p < .001$, indicating the influence of the body is smaller at the extreme ends of the face continuum.

We calculated the median reaction times by participant and condition and performed the same 2×5 ANOVA on the reaction time data. This revealed no significant effects, indicating the results are not biased by a speed–accuracy trade-off.

Discussion

The results of Experiment 2 provide clear evidence that recognition of facial expressions is influenced by the accompanying body language. A happy face on a happy body is categorized more frequently as happy, compared to when the same happy face appears on a fearful body. And a fearful face on a fearful body is categorized as more fearful, compared to when it appears in combination with a happy body expression. It should be stressed that the instructions explicitly stated to categorize the facial expression, so there was no ambiguity regarding the “target” for classification.

These results are consistent with a previous study using compound stimuli of angry and fearful facial and whole-body expressions (Meeren et al., 2005), whose findings are now extended to the emotions fear and happiness.

Moreover, the interaction and trend analysis reported in this study indicate that the influence of the body expression is a function of the ambiguity of the facial expression: the whole-body expression has the most influence when the face ambiguity is highest and decreases with reduced facial ambiguity.

Experiment 3: The Influence of Body Language on Recognition of Voice Prosody

Multisensory integration is considered adaptive, because it reduces stimulus ambiguity (de Gelder & Bertelson, 2003). Previous studies have indicated that facial expressions and emotional tone of voice or emotional prosody influence each other (de Gelder & Vroomen, 2000; Massaro & Egan, 1996). Emotional prosody refers to the variations in melody, intonations, pauses, stresses, and accents of speech. Factors that play a role in voice prosody of some

emotions are duration and intonation. For example, a happy sentence is of normal duration, the pitch is high, and there is a major change in pitch. Acoustically, “fear” is very similar to happiness and the duration is also normal, the mean pitch is also high, but the change in pitch is smaller than in a happy sentence. Integration of affective information from different sensory channels seems to be essential for accurate and fast recognition of emotions.

Developmental studies on recognition of prosody typically present facial expressions with either prosodic congruent or incongruent vocal expressions, while measuring the looking time at the faces (Soken & Pick, 1992; Walker, 1982; Walker & Grolnick, 1983; Walker-Andrews, 1986). The results indicate that infants can already detect changes in prosody at 3 months of age (Walker & Grolnick, 1983).

In a study with static facial expressions and emotional spoken sentences, de Gelder and Vroomen (2000) observed a cross-modal influence of the affective information. Recognition of morphed vocal expressions was biased toward the simultaneously presented facial expression, even when the participants were instructed to ignore the visual stimuli. A follow up study suggests that this cross-modal integration of affective information takes place automatically, independent of attentional factors (Vroomen, Driver, & de Gelder, 2001) and works also when the observer is unaware of the expression of the face, as observed in a cortically blind patient (de Gelder et al., 2005). Investigations of the time course of this integration with ERP have indicated that affective information from different sensory channels is combined early in the perceptual process (de Gelder et al., 1999). In Experiment 3, we used a similar paradigm as de Gelder and Vroomen (2000) (Experiment 3), but we tested for the effect of whole-body expressions instead of facial expressions.

Method

Participants. The group consisted of 16 neurologically intact participants (mean age = 32.9 years).

Materials and procedure. The visual stimuli were the same fearful and happy whole-body expressions as in Experiment 2, but with the faces blurred (see Figure 3A for an example). The auditory stimulus materials consist of a spoken sentence, edited as to express different levels of emotion on a 7-step continuum between fearful and happy. The editing consisted of adjusting the duration, pitch range and pitch register (see de Gelder & Vroomen, 2000, for details). For the present study, we only used the last four words of the sentence (“kwam met het vliegtuig,” meaning “arrived by plane”). These lasted about 600 ms.

Audiovisual stimuli were created by pairing each body expression with each of the seven auditory stimuli, thus resulting in 14 audiovisual stimuli: the seven vocal expressions paired with the fearful body and the same seven vocal expressions paired with the happy body.

The visual stimuli were presented on a computer screen. Auditory stimuli were presented at a comfortable listening level over loudspeakers. The on- and off-set of the visual stimulus was synchronized with the auditory stimulus. Participants had a maximum of 4,000 ms to respond, followed by an intertrial interval of 1,000 ms.

The task was to categorize the expression of the voice (fearful or happy) in a two-alternative forced choice task. On catch trials, a

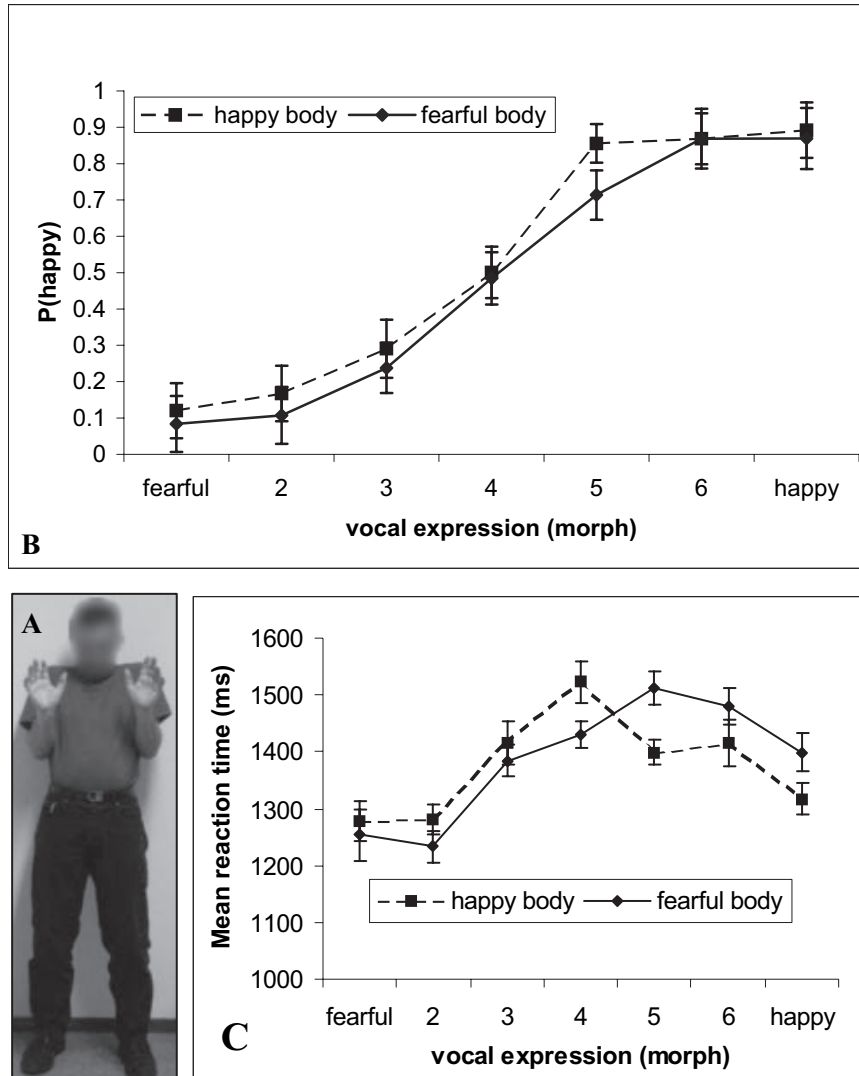


Figure 3. (A) Fearful whole-body expression as presented in Experiment 3. (B) Mean proportion happy responses in Experiment 3 as a function of vocal expression. (C) Mean reaction time in Experiment 3 as a function of vocal expression. Error bars represent 1 SEM around the mean.

white “X” appeared on the body, and participants were to refrain from responding. We included these catch trials to make sure participants perceived the body. The experiment was run in two sessions each with both 49 randomized trials (3 presentations of all 14 audiovisual stimuli + 7 catch trials). The sessions were preceded by 10 practice trials.

Results

Only the participants that missed no more than five catch trials were selected for the analysis. For this reason, two participants were excluded from the analysis. The proportion of *happy* responses was calculated for each participant for every combination of voice prosody and body expression on the experimental trials. Results are displayed in Figure 3B.

The data were submitted to a repeated measures ANOVA with voice (7 levels) and body (2 levels) as within-subjects factors. This

showed a main effect of voice, $F(6, 78) = 24.90, p < .001$, and body, $F(1, 13) = 9.94, p < .008$, but no interaction.

We calculated the median reaction times by participant and condition. Reaction time data are shown in Figure 3C. The 7×2 repeated measures ANOVA on the reaction time data only showed a main effect of voice, $F(6, 78) = 9.43, p < .001$. This simply reflects the fact that reaction times increase as the vocal expression becomes more ambiguous, as can be seen in Figure 3C.

Discussion

The results of Experiment 3 indicate an influence of a perceived whole-body expression on the recognition of voice prosody. When observers make judgments about the emotion conveyed in a voice, recognition is biased toward the simultaneously perceived whole-body expression. The task required attention to be focused exclu-

sively to the voice, but nevertheless, there is a systematic influence of body expression.

The cross-modal affective bias effect has also been observed between voice prosody and facial expressions and seems to be mandatory and automatic (de Gelder & Vroomen, 2000; Vroomen et al., 2001). Vroomen et al. (2001) investigated whether bimodal integration of affective faces and voices required limited attentional resources. Subjects judged whether a voice expressed happiness or fear, while instructed to ignore a concurrently presented static facial expression. Additional tasks were added, to manipulate the attentional load. The cross-modal bias effect was independent of whether the subjects performed a demanding attentional task. In the same line, the present experiment suggests perceptual integration of bimodal emotion expression rather than integration of the two sources based on a later post-perceptual and more cognitive process as suggested by previous literature process. ERP studies with audiovisual affective stimuli point to an early integration of sensory modalities (within 110 ms poststimulus onset), also suggesting a perceptual mechanism, instead of a later more cognitive process (Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000). A study with intracranial recordings in monkeys indicated integration of facial and vocal signals in primary auditory cortex through enhancement and suppression of field potentials (Ghazanfar, Maier, Hoffman, & Logothetis, 2005). Most important, the combination of a fearful face with a fearful tone of voice increases activation in amygdala (Dolan, Morris, & de Gelder, 2001), indicating that the merging of information across stimulus categories is driven by the perception of the meaning irrespective of the medium through which the meaning is conveyed. These questions need to be addressed in follow-up studies using methods that provide a better insight in the temporal dynamics.

General Discussion

Experiment 1 showed that the newly developed body stimuli are easily recognized when no verbal labels are provided. We also observed fear to be the most difficult emotion to be recognized, consistent with previous reports on facial expressions (Milders et al., 2003).

In Experiment 2, we found perception of facial expressions to be biased toward the not explicitly attended and task irrelevant body language. This replicates our findings of a previous study (Meeren et al., 2005), but also extends the observations to other combinations of emotions and indicates that the magnitude of the influence of the body expression depends on facial expression ambiguity. The data of Experiment 3 show that when participants are asked to identify the emotional tone of a voice, while ignoring a simultaneously presented body, they are nevertheless susceptible to be influenced by the bodily expression.

The results of the present study clearly indicate the importance of whole-body expressions as significant emotional stimuli and reveals similarities with findings from facial expression research. The presence of an unattended expressive body influences recognition of faces and auditory stimuli.

From an evolutionary perspective, an important adaptive function of body language is communication of relevant information to other members of the species. Especially in social species, there are considerable adaptive benefits in the ability to interpret emo-

tional displays by conspecifics (Dawkins & Krebs, 1978). This is especially the case when the facial expression of the producer is not visible, for example, because of the viewers' perspective or because of a too great distance to the source.

In the case of emotional body language, perceiving dynamics seem to be particularly important. Recognition of dynamic whole-body expressions is easier than static stimuli (Atkinson et al., 2004) and seems little affected by cultural factors (Hejmadi, Davidson, & Rozin, 2000; Rozin, Taylor, Ross, Bennett, & Hejmadi, 2005). The present studies used static images, in line with the large majority of studies of facial expressions, but there is reason to believe that the important dynamic emotion information may not need to be present explicitly to create a dynamic percept. When viewing two successive presentations of a stimulus object with implied motion, subjects fail to notice the difference between them if the second one represents the same event, but a moment later in time (Freyd, 1983). Moreover, viewing implied motion stimuli activates brain area MT/MST, involved in the processing of movement (Kourtzi & Kanwisher, 2000). For the case of emotional whole-body expressions, we observed that viewing static fear images yield strong activity in motor areas (de Gelder et al., 2004). More recently, we compared activation for static versus dynamic presentation of the same images and observed no difference in amygdala activity for the two presentation conditions (Grèzes et al., 2006). Thus, there is reason to believe that the sight of a bodily expression of emotion affects the viewer profoundly even when motion is not explicitly shown. Creating this emotional movement illusion is indeed what the visual arts have excelled at for a very long time.

References

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, 12(2), 169–177.
- Argyle, M. (1988). *Bodily communication*. London: Methuen.
- Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature Neuroscience*, 7, 542–548.
- 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, 717–746.
- Benson, P. J., & Perrett, D. (1991). Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images. *European Journal of Cognitive Psychology*, 3, 105–135.
- Bertenthal, B. I., Proffitt, D. R., & Kramer, S. J. (1987). Perception of biomechanical motions by infants: Implementation of various processing constraints. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 577–585.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.
- Dawkins, R., & Krebs, J. R. (1978). Animals' signals: Information or manipulation. In J. R. Krebs & N. B. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (pp. 282–309). Oxford: Blackwell Scientific Publications.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7, 242–249.
- de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, 7, 460–467.
- de Gelder, B., Bocker, K. B., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: Early interaction revealed by human electric brain responses. *Neuroscience Letters*, 260, 133–136.

- de Gelder, B., Morris, J. S., & Dolan, R. J. (2005). Unconscious fear influences emotional awareness of faces and voices. *Proceedings of the National Academy of Sciences of the USA*, *102*, 18682–18687.
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proceedings of the National Academy of Sciences of the USA*, *101*, 16701–16706.
- de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*, 289–311.
- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Sciences of the USA*, *98*(17), 10006–10010.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*, 2470–2473.
- Ekman, P. (1965). Differential communication of affect by head and body cues. *Journal of Personality and Social Psychology*, *2*, 726–735.
- Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics*, *33*, 575–581.
- Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *Journal of Neuroscience*, *25*, 5004–5012.
- Gliga, T., & Dehaene-Lambertz, G. (2005). Structural encoding of body and face in human infants and adults. *Journal of Cognitive Neuroscience*, *17*, 1328–1340.
- Grèzes, J., Pichon, S., & de Gelder, B. (2006). Perceiving fear in dynamic body expressions. *Neuroimage*, *35*, 959–967.
- Grosbras, M. H., & Paus, T. (2006). Brain networks involved in viewing angry hands or faces. *Cerebral Cortex*, *16*, 1087–1096.
- Hadjikhani, N., & de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Current Biology*, *13*, 2201–2205.
- Hejmadi, A., Davidson, R. J., & Rozin, P. (2000). Exploring Hindu Indian emotion expressions: Evidence for accurate recognition by Americans and Indians. *Psychological Science*, *11*, 183–187.
- Kourtzi, Z., & Kanwisher, N. (2000). Activation in human mt/mst by static images with implied motion. *Journal of Cognitive Neuroscience*, *12*, 48–55.
- Massaro, D. W., & Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin and Review*, *3*, 215–221.
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the USA*, *102*, 16518–16523.
- Milders, M., Crawford, J. R., Lamb, A., & Simpson, S. A. (2003). Differential deficits in expression recognition in gene-carriers and patients with Huntington's disease. *Neuropsychologia*, *41*, 1484–1492.
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, *93*, 603–608.
- Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, *11*, 1329–1333.
- Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science*, *14*, 302–308.
- Rozin, P., Taylor, C., Ross, L., Bennett, G., & Hejmadi, A. (2005). General and specific abilities to recognise negative emotions, especially disgust, as portrayed in the face and the body. *Cognition and Emotion*, *19*, 397–412.
- Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *Journal of Neuroscience*, *25*, 11055–11059.
- Slaughter, V., Stone, V. E., & Reed, C. (2004). Perception of faces and bodies: Similar or different? *Current Directions in Psychological Science*, *13*, 219–223.
- Soken, N. H., & Pick, A. D. (1992). Intermodal perception of happy and angry expressive behaviors by seven-month-old infants. *Child Development*, *63*, 787–795.
- Sprengelmeyer, R., Young, A. W., Schroeder, U., Grossenbacher, P. G., Federlein, J., Buttner, T., et al. (1999). Knowing no fear. *Proceedings of the Royal Society B: Biological Sciences*, *266*, 2451–2456.
- Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: An erp study on the body-inversion effect. *Neuroreport*, *15*, 777–780.
- Vroomen, J., Driver, J., & de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? *Cognitive, Affective & Behavioral Neuroscience*, *1*, 382–387.
- Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology*, *33*, 514–535.
- Walker, A. S., & Grolnick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behavior and Development*, 491–498.
- Walker-Andrews, A. S. (1986). Intermodal perception of expressive behaviors: Relation of eye and voice? *Developmental Psychology*, *22*, 373–377.

Received June 30, 2006

Revision received January 3, 2007

Accepted January 29, 2007 ■